ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

The Cost of Capital and Misallocation in the United States

Authors	Miguel Faria e Castro, Julian Kozlowski, and Jeremy Majerovitz
Working Paper Number	2025-013B
Revision Date	June 2025
Citable Link	https://doi.org/10.20955/wp.2025.013
Suggested Citation	Faria e Castro, M., Kozlowski, J., Majerovitz, J., 2025; The Cost of Capital and Misallocation in the United States, Federal Reserve Bank of St. Louis Working Paper 2025-013. URL https://doi.org/10.20955/wp.2025.013

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The Cost of Capital and Misallocation in the United States*

Miguel Faria-e-Castro — FRB of St. Louis
Julian Kozlowski — FRB of St. Louis
Jeremy Majerovitz — University of Notre Dame

July 14, 2026

Abstract

We propose a framework to estimate capital misallocation using credit registry data and apply it to the United States. A dynamic corporate finance model maps loan-level observables into three rates: the lender's discount rate, the firm's private cost of capital, and the social cost of capital. We derive a sufficient statistic for misallocation that depends only on the mean and dispersion of the social cost of capital. Applying the framework to U.S. loan originations, we find substantial heterogeneity in the cost of capital but modest aggregate losses in normal times: reallocating capital across firms would raise output by about 0.7%. These losses more than doubled during 2020–2021, rising to 1.6%. The increase was driven primarily by greater dispersion in lender discount rates, reflecting both weaker pass-through from expected losses to loan interest rates and a rise in the cross-sectional variance of expected losses.

JEL Classification: D24, E22, E44, O47, O51.

Keywords: cost of capital, misallocation, credit registry data.

*We thank our discussant Vladimir Yankov as well as Juan Sanchez, François Gourio, and participants in seminars and conferences for thoughtful questions and feedback. We thank Zeina Shalaby and Nan Jiang for excellent research assistance. Majerovitz thanks the University of Notre Dame's Building Inclusive Growth (BIG) Lab for support. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis, the Board of Governors of the Federal Reserve, or the Federal Reserve System.

1 Introduction

How much does it cost a firm to obtain capital? Economic models often simplify by assuming that all firms can borrow in a competitive market at a common rate. In reality, however, the cost of capital varies significantly across firms. This variation stems not only from differences in contractual interest rates but also from firm-specific factors such as default probabilities, loan terms, and lender cost of funds. Such heterogeneity in the cost of capital has profound implications: it can distort the allocation of capital across firms, leading to inefficiencies in economic output (Gilchrist et al., 2013; Hsieh and Klenow, 2009). Understanding these inefficiencies is critical for policymakers and researchers seeking to design effective financial and economic policies.

This paper makes two contributions to the literature. First, we develop a novel method that leverages a dynamic corporate finance model and uses credit registry microdata to measure the dispersion of the cost of capital. The central insight is that in equilibrium the firms equate the cost of capital to the marginal product of capital. This allow us to use credit registry data to measure dispersion in the cost of capital and quantify the dispersion in the marginal product of capital and therefore capital misallocation. The theory implies a sufficient statistic for misallocation that can be directly measured with credit registry data. Second, we apply this method to U.S. data and uncover two primary insights. While the cost of capital is heterogeneous across firms, the implied misallocation is small. This misallocation, however, exhibits a substantial amount of variation over time, and, in particular, it doubled in 2020-21, around the COVID-19 pandemic.

The method we develop offers several advantages. Unlike traditional approaches that require strong assumptions about production functions as well as detailed data on firms financials such as sales or value added, our approach relies on credit registry data that is commonly collected by financial regulators around the developed and developing world, and uses sufficient statistics derived directly from moments of the data. This not only simplifies implementation but also provides more robust identification of the sources of misallocation without heavy reliance on calibration assumptions.

Our estimates capture the misallocation arising from dispersion in the cost of capital among firms that borrow from commercial banks. Our paper therefore follows what Restuccia and Rogerson (2017) refer to as the "direct approach": we measure the consequences of a specific source of misallocation rather than attempting to measure the full cost of misallocation. Therefore, our finding of small losses does not necessarily imply that the economy as a whole is close to efficient. Rather, it suggests that if there is substantial misallocation, it must arise primarily from sources other than the commercial bank lending distortions captured by our measure.

Section 2 describes the dynamic corporate finance model that provides the foundation for the derivation of the sufficient statistic for misallocation and our empirical measurement of the cost of capital. The model captures firm-level borrowing, investment, and default decisions in the presence of idiosyncratic shocks, such as productivity fluctuations or stochastic fixed operating costs. Each firm borrows from a single lender who discounts future cash flows at a match-specific rate ρ. While we refer to ρ as the lender's discount rate, it more broadly reflects variation in loan pricing that is not fully explained by observable loan terms. Using the firm's optimality conditions, we show how ρ influences the firm's internal cost of capital, which in turn determines its expected marginal revenue product of capital. This link forms the basis for our measure of misallocation.

In Section 3, we theoretically map heterogeneity in the cost of capital to economic efficiency costs. We compare the decentralized equilibrium to the allocation chosen by a planner who reallocates capital across firms to maximize aggregate output, subject to the same aggregate capital stock and taking firms' default decisions as given. This corresponds to the misallocation of capital on the intensive margin, which is the main form of misallocation studied by the misallocation literature (e.g. Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008). We define the social cost of capital, which is the marginal value of allocating an extra unit of capital for each firm from the point of view of the planner. We show that this measure is approximately equal to the sum of two terms: the lender's discount rate plus a term that reflects financial frictions related to limited liability and recovery in case of default in the spirit of Cooley and Quadrini (2001). At the optimum, the planner would like to equate the social cost of capital across firms. This insight allows us to derive a sufficient statistic for the output loss from misallocation that depends only on the mean and variance of the social cost of capital. This statistic is robust to firm-level heterogeneity in production technologies and does not rely on structural estimation.

Section 4 describes how we use credit registry data for measurement. We define the lender's discount rate as the internal rate of return that satisfies the lender's break-even condition, accounting for both repayment probabilities and expected losses in default. To compute this rate at the firm level, we require loan-level data on contractual interest rates, loan maturities, borrower-specific probabilities of default, and loss given default (LGD), as well as the expected inflation term structure. In the case of floating-rate loans, we also need expectations for forward benchmark interest rates. Using these variables as well as the equations of the model, we estimate three distinct rates: the lender's discount rate, the firm's cost of capital, and the social cost of capital. We then apply the sufficient statistic to the social cost of capital to estimate the output loss from misallocation.

Section 5 presents our empirical findings. Using data from more than 67,000 loans originated since 2014Q4, we show that the average measures of cost of capital closely track the five-year U.S. Treasury rate. We estimate three distinct rates. First, the *lender's discount rate* is specific to each borrower-lender pair; it has a mean of 2.0% and a standard deviation of 1.5% over the period in analysis. Second, the *firm's cost of capital*—defined as the expected payment by the firm—has a mean of 1.0% and a standard deviation of 2.6%. This rate is lower than the lender's discount rate because the firm does not internalize recoveries in the case of default, unlike the lender. Finally, the *social cost of capital* reflects the total return on capital from a social perspective, incorporating expected recovery in the event of default. It has a mean of 1.8% and a standard deviation of 1.7%.

At the optimum, the planner seeks to equalize the social cost of capital across firms. Our sufficient statistic provides a mapping from the dispersion of the social cost of capital to output losses due to misallocation. We estimate that, under normal conditions, the implied output loss from capital misallocation is modest, around 0.7%. However, this loss increased significantly during the COVID-19 pandemic (2020–2021), rising to 1.6%.

We then investigate why misallocation increased in 2020–2021. The first decomposition separates the social cost of capital into the lender discount rate and the financial-frictions term coming from default and recovery. This exercise shows that the increase was driven primarily by greater dispersion in the lender discount rate. We then study why this dispersion rose. The lender discount rate reflects the gap between loan interest rates and expected losses; in an efficient benchmark, higher expected losses should be passed through one-for-one into higher interest rates. In the data, this pass-through was already incomplete before the pandemic and weakened sharply after 2020. We find that approximately 42% of the 2020–2021 increase in misallocation is explained by changes in the relationship between expected losses and the cost of capital, while roughly 41% is explained by the increase in the variance of expected losses. Changes in average expected losses and residual variation not explained by expected losses play much smaller roles. Thus, the pandemic increase reflects not simply higher average credit risk, but a deterioration in how loan prices adjusted to the cross-sectional distribution of risk.

Section 6 reports four extensions. First, we risk-adjust the lender discount rate by removing estimated compensation for exposure to aggregate risk factors. The resulting misallocation series is close to the benchmark and, if anything, shows a larger increase during the pandemic, indicating that aggregate risk premia do not explain the main result. Second, we compare our social cost of capital to standard ARPK-based measures of misallocation. The two are positively correlated, especially for value-added-based ARPK, but ARPK dispersion implies much larger losses than our credit-based sufficient statistic. Third, we compare our lender discount rate to the text-based cost-of-capital measure of Gormsen and Huber (2023) and find a positive relationship that remains after controlling for time and industry effects. Fourth, we place the U.S. estimates in a cross-country comparison using loan-market statistics from Pakistan, Brazil, and Mexico. The comparison suggests that U.S. credit markets exhibit relatively low implied misallocation, while also illustrating how our approach could be applied more broadly with credit registry microdata.

Literature Review. Our paper contributes to the broader literature on measuring misallocation. Following seminal work by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), there has been significant progress in quantifying misallocation across various settings (see Hopenhayn, 2014 and Restuccia and Rogerson, 2017 for comprehensive reviews). A key challenge in this literature is measuring misallocation without imposing strong assumptions on firms' production technologies. Haltiwanger et al. (2018) emphasize that standard approaches are only valid under restrictive assumptions, such as a common Cobb-Douglas production function with firm-specific productivity shifters.

One strand of the literature focuses on specific sources of distortions. For instance, Kaymak and Schott (2024) study corporate tax asymmetries and find that heterogeneity in effective marginal tax rates can distort capital and labor allocation, reducing aggregate productivity. Alternatively, recent work has sought to directly estimate marginal products using (quasi-)experimental variation, allowing for richer production heterogeneity (e.g. Carrillo et al., 2023; Hughes and Majerovitz, 2025). However, such approaches have only been applied in narrow contexts where experimental variation is available.

Our paper measures heterogeneity in the marginal product of capital by exploiting firm-level variation in the cost of capital, allowing us to assess misallocation across a much broader set of firms while remaining agnostic to functional form assumptions. Closest to our approach, Gilchrist et al. (2013) develop a tractable framework to quantify misallocation arising from dispersion in borrowing costs. We emphasize two key differences. First, our framework explicitly models default, whereas Gilchrist et al. (2013) control for default risk in a reduced-form way, by regressing credit spreads on distance to default. Second, their analysis relies on corporate bond data, which restricts attention to large firms. In contrast, we use bank loan data that encompasses a significantly wider range of firms, including small, medium, and large-sized enterprises.

We also contribute to a literature that estimates heterogeneity across firms in interest rates and/or the cost of capital. Banerjee and Duflo (2005) summarize early evidence for substantial heterogeneity in interest rates across borrowers in developing countries, arguing that this heterogeneity implies significant misallocation. Recent work by Gormsen and Huber (2023, 2024) analyzes transcripts of firm earnings calls to extract information on the discount rates and cost of capital that firms use. Cavalcanti et al. (2024) use credit registry data to study heterogeneity in interest rates for borrowing firms in Brazil. They find substantial heterogeneity across firms and use a dynamic structural model with financial frictions to infer the cost of capital. This paper also builds on the findings of Faria-e-Castro et al. (2024), who analyze the dispersion in borrowing rates for U.S. firms using a comprehensive database of loans and bonds. Their study highlights significant heterogeneity in borrowing costs, even within firms, and demonstrates the persistent impact of borrowing costs on firm-level investment and borrowing behaviors.

Relative to this previous literature, our paper makes two key methodological contributions. First, we provide a method to estimate a firm's cost of capital from credit registry data. This is not as simple as measuring the interest rate because the cost of capital depends on the ex-ante repayment probability and expected losses given default. Second, we show how to use moments of the distribution of the cost of capital to develop sufficient statistics that allow us to measure the cost of misallocation non-parametrically in a dynamic, stochastic model.

2 Corporate Finance Model

This section outlines the core components of the model, which we use as a measurement device. We demonstrate how the model's optimality conditions can be derived and later integrated with microdata on loan characteristics to estimate lender discount rates and the firms' cost of capital. These rates are used to infer the firm's expected marginal product of capital, which is a key input for our measure of misallocation.

The model is set in infinite-horizon discrete time, with periods indexed by t. The economy is populated by firms that borrow and invest, and by lenders who finance those firms. Each firm is matched with a lender. Firms are indexed by i. There is no aggregate risk. We now describe the decision problem of the borrowing firm, and its interaction with the lenders.

Borrowers. The borrowers in the model are firms operating in the nonfinancial sector. These firms operate under limited liability and make decisions regarding production, investment, and borrowing. Output (net of non-capital costs) is generated using a production function $f(k_i, z_i)$, where k_i represents capital and z_i denotes a vector of shocks that affect firm net output. We allow z_i to be a vector; this accommodates productivity shocks, stochastic fixed costs, as well as rich heterogeneity in the production function. To sustain or expand their operations, firms invest in capital k_i and issue long-term defaultable debt b_i.

Timing. At the start of each period, the firm's state is (k_i, b_i, z_i). First, the firm chooses to either default or repay. If the firm defaults, its capital is liquidated and used to repay the lender; the lender receives $\phi_i(k_i)$. The defaulting firm receives zero payoff and exits forever. If the firm repays, it continues to operate, earns profits, and makes scheduled debt payments. Continuing firms must then choose next period's capital and debt, (k_i', b_i'), without yet knowing the next period's draw of z_i'.

Firm's Problem. The value of repayment for a firm is expressed as:

$$V_i(k_i, b_i, z_i) = \max_{k_i', b_i'} \pi(k_i, b_i, z_i, k_i', b_i') + \beta \mathbb{E}\left[\max\left\{V_i(k_i', b_i', z_i'), 0\right\} | z_i\right], \tag{1}$$

where $\pi(k_i, b_i, z_i, k_i', b_i')$ denotes the firm's profit function, and β represents the discount factor. The profit function captures the firm's net return from production and financing decisions:

$$\pi(k_i, b_i, z_i, k_i', b_i') = f(k_i, z_i) + (1-\delta)k_i - k_i' - \theta_i b_i + Q_i(k_i', b_i', z_i)[b_i' - (1-\theta_i)b_i].$$

Here, $f(k_i, z_i)$ represents the firm's net output as a function of capital k_i and idiosyncratic shocks z_i, $(1-\delta)k_i$ accounts for the depreciated value of current capital, and k_i' denotes new capital investment. We model long-term debt as a geometrically decaying perpetuity with rate θ_i. Thus, $\theta_i b_i$ reflects repayment on existing debt, while $Q_i(k_i', b_i', z_i)$ captures the price of newly issued debt, with $b_i' - (1-\theta_i)b_i$ representing net new borrowing.

Lenders. Lenders finance firms, with each firm matched to a single lender. Upon matching, the borrower-lender pair draws a realization of ρ_i, which represents the discount rate that the lender uses to price debt. [1] For this reason, we refer to ρ_i as the lender's discount rate. Loans are priced so that lenders break even using ρ_i as their discount rate, taking into account firm-specific characteristics and risk assessments.

Debt Pricing. Lenders are risk-neutral and price debt based on their discount rate, ρ_i. The price of debt $Q_i(k_i', b_i', z_i)$ is determined as:

$$Q_i(k_i', b_i', z_i) = \frac{\mathbb{E}\left[\mathcal{P}_i(k_i', b_i', z_i')\left[\theta_i + (1-\theta_i)Q_i(k_i'', b_i'', z_i')\right] + (1-\mathcal{P}_i(k_i', b_i', z_i'))\,\frac{\phi_i(k_i')}{b_i'}\,\middle|\, k_i', b_i', z_i\right]}{1+\rho_i}, \tag{2}$$

where $\mathcal{P}_i(k_i', b_i', z_i')$ is an indicator function that is equal to 1 if the firm repays, and 0 otherwise, and $\phi_i(k_i')/b_i'$ is the recovery rate in the event of default, per dollar lent.

Solution to the Firm's Problem. The solution to the firm's problem in (1) is characterized by two first-order conditions, with respect to capital k_i' and debt b_i':[2]

$$0 = -1 + \frac{\partial Q_i(k_i', b_i', z_i)}{\partial k_i'}[b_i' - (1-\theta_i)b_i] + \beta\mathbb{E}\left[\mathcal{P}_i(k_i', b_i', z_i')[f_k(k_i', z_i') + 1 - \delta]\middle|\, z_i\right]$$

$$0 = \frac{\partial Q_i(k_i', b_i', z_i)}{\partial b_i'}[b_i' - (1-\theta_i)b_i] + Q_i(k_i', b_i', z_i) - \beta\mathbb{E}\left[\mathcal{P}_i(k_i', b_i', z_i')[\theta_i + (1-\theta_i)Q_i(k_i'', b_i'', z_i')]\middle|\, k_i', b_i', z_i\right]$$

[1] This variable captures both lender- and borrower-specific factors that lie outside the scope of the model, such as lender financing costs, risk appetite, or the dynamics of relationship lending. We do not provide a specific microfoundation for the heterogeneity in ρ_i, and instead focus on analyzing its implications.

[2] We provide a derivation of this equation in Appendix A.

where we use the envelope conditions $\frac{\partial V_i}{\partial k_i} = f_k(k_i, z_i) + 1 - \delta$ and $\frac{\partial V_i}{\partial b_i} = -[\theta_i + (1-\theta_i)Q_i]$. Throughout, we assume that the firm chooses an interior solution for debt. Simplifying notation, so that $\mathcal{P}'_i$ is a shorthand for $\mathcal{P}_i(k'_i, b'_i, z'_i)$ and Q'_i for $Q_i(k''_i, b''_i, z'_i)$, we can combine the two first-order conditions to write:

$$
\begin{aligned}
\mathbb{E}\left[\mathcal{P}'_i \cdot (f_k(k'_i, z'_i) + 1 - \delta) \middle| z_i\right] &= \frac{\mathbb{E}\left[\mathcal{P}'_i \cdot (\theta_i + (1-\theta_i)Q'_i) \middle| k'_i, b'_i, z_i\right]}{Q_i} \cdot \frac{1 - \frac{\partial Q_i}{\partial k'_i}[b'_i - (1-\theta_i)b_i]}{1 + \frac{\partial Q_i}{\partial b'_i}\frac{[b'_i - (1-\theta_i)b_i]}{Q_i}} \\
&= (1 + r_i^{firm}) \cdot \mathcal{M}_i \qquad (3)
\end{aligned}
$$

This equation tells us that the firm sets its expected marginal product of capital equal to its perceived marginal cost of funds $1 + r_i^{firm}$, multiplied by an adjustment factor $\mathcal{M}_i$ that reflects the impact of its decisions on the price of debt. We discuss these terms further below.

Firm's Cost of Capital. We define the firm's cost of capital, $1 + r_i^{firm}$, as the ratio of the expected value of future repayments adjusted for the probability of repayment, $\mathbb{E}\left[\mathcal{P}'_i \cdot (\theta_i + (1-\theta_i)Q'_i)\right]$, to the current price of borrowing, Q_i. The firm's cost of capital is the expected implicit interest rate that it pays on its debt. Formally, we define it as:

$$
1 + r_i^{firm} := \frac{\mathbb{E}\left[\mathcal{P}'_i \cdot (\theta_i + (1-\theta_i)Q'_i) \middle| k'_i, b'_i, z_i\right]}{Q_i}. \qquad (4)
$$

We can also relate the firm's cost of capital to the lender's discount rate ρ_i. Combining the definition above with Equation (2), we obtain:

$$
1 + r_i^{firm} = (1 + \rho_i) - \underbrace{\mathbb{E}\left[1 - \mathcal{P}'_i \middle| k'_i, b'_i, z_i\right] \cdot \frac{\phi_i(k'_i)}{Q_i b'_i}}_{\text{Expected Recoveries}} \qquad (5)
$$

The firm's cost of capital differs from the lender's discount rate due to recovery in the event of default. When there is no recovery ($\phi_i = 0$), the firm's cost of capital equals the lender's discount rate ($r_i^{firm} = \rho_i$). On the other hand, when the lender can recover some value after default ($\phi_i > 0$), the firm's cost of capital r_i^{firm} is lower than ρ_i. This reduction in perceived borrowing cost occurs because the borrower only accounts for states where repayment occurs. Thus the firm expects to pay a lower rate than the lender expects to recover.

Price Feedback Multiplier. The price feedback multiplier $\mathcal{M}_i$ is the second term in Equation (3). It captures the effect of the firm's borrowing and investment on the price of debt. We can rewrite it as:

$$\mathcal{M}_i := \frac{1 - \gamma_i \times \frac{Q_i \cdot b_i'}{k_i'} \times \frac{\partial \log Q_i}{\partial \log k_i'}}{1 + \gamma_i \times \frac{\partial \log Q_i}{\partial \log b_i'}}, \qquad \gamma_i := \frac{b_i' - (1-\theta_i) b_i}{b_i'},$$

where γ_i measures the share of tomorrow's debt that will be newly purchased. The numerator of $\mathcal{M}_i$ reflects the feedback from changes in capital on the price of debt, while the denominator incorporates the feedback from changes in borrowing. Together, these terms capture how endogenous prices influence the firm's cost of capital.

3 Measuring Misallocation

We now turn from the individual firm's problem to the aggregate economy. The previous section shows how financing conditions affect a firm's investment choice and expected marginal product of capital. In this section, we ask how dispersion in these firm-level costs of capital translates into aggregate output losses. To do so, we define aggregate output and welfare, characterize the allocation chosen by a planner facing the same information structure as firms, and isolate the intensive-margin reallocation of capital across firms. This exercise identifies the relevant object for measuring misallocation: the social cost of capital, which captures the marginal value of allocating an additional unit of capital to each firm from the planner's perspective. We then derive a sufficient statistic that maps the mean and dispersion of the social cost of capital into the output loss from capital misallocation.

3.1 The Aggregate Economy and Welfare

We begin by setting up the aggregate environment in order to study both the decentralized equilibrium and the planner's problem. The firm's problem is the same as before. There is no aggregate risk, so aggregates are not stochastic. Firms make undifferentiated products and take the price of their output as given.

Entry and Exit. We use N_t to denote the set of firms that can operate at time t. Each period, a set of new firms E_t enters the economy. New entrants start with zero capital and do not produce in period t. The set E_t is thus not included in N_t, but is included in N_{t+1}. We use X_t to denote the firms that exit in period t. The set X_t is included in N_t but is not included in N_{t+1}. The law of motion for N_t is:

$$N_{t+1} = (N_t \setminus X_t) \cup E_t \tag{6}$$

There is an aggregate resource cost to entry, which we denote C_t^{entry}. This cost is a function of the set of entering firms:

$$C_t^{\text{entry}} = C^{\text{entry}} (E_t) \tag{7}$$

Aggregate Output. We compute aggregate output in each period by aggregating across the set of firms in N_t. As in the previous section, the indicator $\mathcal{P}_{i,t}$ will be equal to zero if the firm i exits in period t, and one otherwise. Aggregate output is given by:

$$Y_t = \int_{i \in N_t} \underbrace{\mathcal{P}_{i,t} \cdot f (k_{i,t}, z_{i,t})}_{\text{Output if Operates}} - \underbrace{(1 - \mathcal{P}_{i,t}) \cdot ((1 - \delta) \, k_{i,t} - \phi_i (k_{i,t}))}_{\text{Losses if Defaults}} \, di \tag{8}$$

Note that we have defined output, Y_t, so that it includes both the firm's output in the event of production, $f (k_{i,t}, z_{i,t})$, and the losses from liquidation, $(1 - \delta) \, k_{i,t} - \phi_i (k_{i,t})$, in the event of default. This will allow us to define aggregate investment simply.

Aggregate Capital. There is some initial stock of capital K_0, and future capital depends on investment and depreciation through the standard law of motion. We define aggregate investment in the standard way:

$$I_t = K_{t+1} - (1 - \delta) \, K_t \tag{9}$$

Finally, aggregate capital is given by:

$$K_t = \int_{i \in N_t} k_{i,t} di \tag{10}$$

Welfare. There is a representative household that obtains utility from consumption: we abstract from inequality to focus on productive efficiency. The household's utility is additively separable over time. Consumption is equal to aggregate output minus investment and entry costs. Thus, welfare in this economy is given by:

$$U = \sum_{t=0}^{\infty} \beta^t \cdot u\left(Y_t - I_t - C_t^{\text{entry}}\right)$$

where β is the household's discount rate and $u(\cdot)$ is the utility it gets from consumption.

3.2 The Planner's Problem

The planner wishes to maximize welfare, U, by choosing each firm's capital, exit, and entry. However, the planner is subject to the same information constraints as the firm. The planner must decide $k_{i,t}$ in period $t-1$, without yet knowing $z_{i,t}$. Exit decisions are made after $z_{i,t}$ is revealed, but with the values for future periods still unknown.

Let $S_i^t := \{z_{i,s}\}_{s=0}^{t}$ denote the entire history of states, through period t.[3] Define $S^t := \{S_i^t\}_{\forall i}$ as the collection of all firms' histories. We can use this notation to set up the appropriate constraints to the planner's problem. The planner must set entry (E_t), exit (X_t), and tomorrow's capital ($k_{i,t+1}$), as functions of the state history S^t.[4] The planner's problem is:

$$U^* = \max_{\left\{E_t(S^t), X_t(S^t), \{k_{i,t+1}(S^t)\}_{i \in N_{t+1}}\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \cdot u\left(Y_t - I_t - C_t^{\text{entry}}\right)$$

s.t.

Equations (6), (7), (8), (9), and (10) hold

[3] Note that the only shock in our model is $z_{i,t}$, so this is the full history of states.

[4] In practice, since there is no aggregate risk, the planner will only need to use the individual firm's state histories to make decisions.

Note that in period $t = 0$, there is an initial set of firms N_0, and initial capital $k_{i,0}$ is set exogenously at those firms.

We can rewrite the planner's problem as a nested maximization problem, to isolate the intensive-margin choice of capital, holding aggregate capital and the extensive margin fixed. Note that $I_t = K_{t+1} - (1-\delta) K_t$, and so it depends only on aggregate capital (not the allocation across firms). Similarly, the entry cost C_t^{entry} depends solely on E_t. We can thus rewrite the planner's problem in the following nested form:

$$U^* = \max_{\{E_t(S^t), X_t(S^t), K_{t+1}(S^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \cdot u\left(\left(\max_{\{k_{i,t}(S^{t-1})\}_{i \in N_t}} Y_t\right) - I_t - C_t^{\text{entry}}\right)$$

subject to the same constraints as before.[5]

3.3 Misallocation of Capital on the Intensive Margin

We can now turn our attention to the inner problem. Note that the inner problem is separable across time periods, allowing us to separate it into a sequence of static problems. We focus on the cost of misallocation in terms of output. Simplifying our notation, we can rewrite the problem as follows:

$$Y_t^*(K_t, N_t) = \max_{\{k_{i,t}\}_{i \in N_t}} \int_{i \in N_t} \mathbb{E}_{t-1}\left[\mathcal{P}_{i,t} \cdot f(k_{i,t}, z_{i,t}) - (1 - \mathcal{P}_{i,t}) \cdot ((1-\delta) k_{i,t} - \phi_i(k_{i,t}))\right] di$$

$$\text{s.t.}$$

$$K_t = \int_{i \in N_t} k_{i,t} di$$

This corresponds to minimizing the misallocation of capital on the intensive margin, which is the main form of misallocation studied by the misallocation literature (e.g. Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008). This problem is now a special case of the environment in Hughes and Majerovitz (2025). We can use their main proposition to derive the cost of misallocation, up to a second-order approximation. Define

$$y_i(k_i) := \mathbb{E}_{t-1}\left[\mathcal{P}_{i,t} \cdot f(k_{i,t}, z_{i,t}) - (1 - \mathcal{P}_{i,t}) \cdot ((1-\delta) k_{i,t} - \phi_i(k_{i,t}))\right].$$

[5]Note that for $t = 0$, the choice of capital is degenerate, since capital is predetermined in that period.

Proposition 1 shows the cost of intensive-margin misallocation.

Proposition 1 (Special Case of Hughes and Majerovitz, 2025)**.** *The cost of intensive-margin misallocation is given by*

$$\underbrace{\log Y_t^*\left(K_t, N_t\right) - \log Y_t}_{\textit{Cost of Intensive-Margin Misallocation}} \approx \frac{1}{2} \cdot \underbrace{\mathbb{E}_{y_i(k_i)}\left[\mathcal{E}_i\right]}_{\textit{Sales-Weighted Elasticity}} \cdot \underbrace{Var_{y_i(k_i)\mathcal{E}_i}\left(\log\left(\frac{\partial}{\partial k_{i,t}} y_i\left(k_i\right)\right)\right)}_{\textit{Weighted Variance of Log Expected MPK}}$$

where $\mathcal{E}_i = -\frac{\left(\frac{\partial}{\partial k_i} y_i(k_i)\right)^2}{y_i(k_i)\cdot \frac{\partial^2}{(\partial k_i)^2} y_i(k_i)}$ *is the elasticity of* $y_i\left(k_i\right)$ *with respect to the cost of capital. Subscripts denote the weights for the weighted moments. All moments are computed for the firms in* N_t*.*

Note that although the proposition above provides a second-order approximation, it becomes exact in a setting where production is Cobb-Douglas and where productivity and distortions are jointly log-normal (the weights also fall out in that special case).

3.4 The Social Cost of Capital

We have already introduced the notion of the lender's discount rate, ρ, and the firm's cost of capital, r^{firm}. We now introduce the notion of the social cost of capital, r^{social}. This will reflect the social marginal product of capital at firm i. We define $1 + r_{i,t}^{social}$ as the derivative of aggregate consumption $(Y_t - I_t - C_t^{\text{entry}})$ at time $t+1$ with respect to k_{it+1}, taking expectations at time t (when the investment decision is made).[6] We have:

$$\begin{aligned} 1 + r_{i,t}^{social} &:= \frac{\partial \mathbb{E}_t\left[Y_{t+1} - I_{t+1} - C_{t+1}^{\text{entry}}\right]}{\partial k_{i,t+1}} \\ &= \mathbb{E}_t\left[\mathcal{P}_{i,t+1}\left(f_k\left(k_{i,t+1}, z_{i,t+1}\right) + 1 - \delta\right) + \left(1 - \mathcal{P}_{i,t+1}\right) \cdot \phi_i'\left(k_{i,t+1}\right)\right] \end{aligned}$$

The social cost of capital for firm i is the shadow value for the planner of allocating an extra unit of capital to that firm. Combining this with the firm's first-order condition for investment in Equation (3) yields:

$$1 + r_{i,t}^{social} = \left(1 + r_{i,t}^{firm}\right) \mathcal{M}_{i,t} + \mathbb{E}_t\left[\left(1 - \mathcal{P}_{i,t+1}\right) \cdot \phi_i'\left(k_{i,t+1}\right)\right] \tag{11}$$

[6] Proposition 1 is in terms of gross output, rather than consumption. Nevertheless, we define r^{social} in this way to parallel our definitions of ρ and r^{firm}.

Note that $1 + r^{social}_{i,t-1} = \frac{\partial}{\partial k_i} y_i(k_i) + 1 - \delta$. This will allow us to use the distribution of r^{social} to measure the cost of misallocation. When we bring this result to the data, we will focus on measuring the variance of r^{social}. Moreover, we will make two further simplifying assumptions. First, we will focus on the unweighted variance, since the weights are difficult to observe in practice. Second, we will use the log-normal approximation $\mathrm{Var}\left(\log\left(r^{social}_{i,t-1} + \delta\right)\right) \approx \log\left(1 + \frac{\mathrm{Var}\left(r^{social}_{i,t-1}+\delta\right)}{\mathbb{E}\left[r^{social}_{i,t-1}+\delta\right]^2}\right)$

To measure misallocation in our data, we combine this with our derivation of r^{social} to yield the following corollary:

Corollary 1. *Assume that $\left(r^{social}_{i,t-1} + \delta\right)$ is log-normally distributed, and also assume that weighted moments can be replaced with unweighted moments. The cost of intensive-margin misallocation is given by*

$$\log Y^*_t\left(K_t, N_t\right) - \log Y_t \approx \frac{1}{2} \cdot \mathcal{E} \cdot \log\left(1 + \frac{Var\left(r^{social}_{i,t-1}\right)}{\mathbb{E}\left[r^{social}_{i,t-1} + \delta\right]^2}\right)$$

This corollary allows us to connect dispersion in r^{social}, an object that we will be able to measure in the microdata, with the cost of intensive-margin misallocation of capital. Intuitively, the planner would like to equalize the shadow value of capital across all firms: this would eliminate dispersion in the social return of capital and result in zero misallocation. We next turn to how to measure ρ, r^{firm}, and r^{social} using credit registry data.

4 Empirical Methodology

This section describes the main data sources that we use, as well as how we map model objects to the data in order to estimate the three rates: the lender discount rate, ρ, the firm's cost of capital, r^{firm}, and the social cost of capital, r^{social}.

4.1 Data Sources

Our main data source is the Schedule H.1 of the FR Y-14Q dataset (Y-14 for short). This is a quarterly regulatory dataset maintained by the Federal Reserve for stress testing purposes, which contains information on individual loan facilities held in the books of the top 30 to 40 bank holding companies (BHCs) in the US. The Y-14 includes all loan facilities exceeding $1 million and we consider data beginning in 2014Q4. Importantly for the purposes of our analysis, the Y-14 contains detailed characteristics of credit facilities such as facility size, origination date and maturity, interest rate or spread, interest rate variability, and the type of loan. Additionally, the Y-14 also covers BHC's risk assessments for each borrower, which include estimates for the 1-year probability of default and loss given default. The probability of default is typically estimated using internal default models that have to be approved by regulators. While there is scope for some discretion in the assignment of these default probabilities (Plosser and Santos, 2018), these models have been subject to standardized guidelines following Basel II (BCBS, 2001). We focus on term loans issued to non-governmental and nonfinancial companies based in the US. Our unit of observation is a loan origination. We do not include credit lines due to lack of information about the fee structure, which would be needed to price these facilities. Appendix B contains a detailed description of the data cleaning procedure and sample restrictions.

In terms of coverage, Faria-e-Castro et al. (2024) show that the FR Y-14Q Schedule H.1 accounts for 91% of Commercial & Industrial lending undertaken by the 25 largest banks in the US (FRED mnemonic: CIBOARD), and 55% of all Commercial & Industrial lending undertaken by all commercial banks in the US (FRED mnemonic: BUSLOANS). Our focus on term loans and relatively stringent cleaning procedures leave us with a total of 67,089 loans.

4.2 Mapping the Model to the Data

An important difference between the model and the data is the payment structure of loans. In the model, for tractability, we assume that firms borrow in long-term debt that is modeled as a perpetuity with geometrically decaying coupons. In the data, on the other hand, we focus our analysis on term loans with a fixed maturity. This section shows how we map model objects to the data, and how we exploit the Y-14 data to retrieve estimates of the lender's discount rate, ρ, the firm's cost of capital, r^{firm}, and the social cost of capital, r^{social}.

Let (i,t) denote a term loan originated at period t. The loan has principal value $B_{i,t}$, maturity $T_{i,t}$, payment schedule $\{D_{i,t,s}\}_{s=1}^{T_{i,t}}$, repayment probability $P_{i,t}$ assumed to be constant over time, and loss given default $LGD_{i,t}$, also constant over time. The break-even condition at origination for a lender with discount rate $\rho_{i,t}$ is given by:

$$B_{i,t} = \mathbb{E}_t \sum_{s=1}^{T_{i,t}} \left[\frac{P_{i,t}^s \cdot D_{i,t,s} + P_{i,t}^{s-1} \cdot (1 - P_{i,t}) \cdot (1 - LGD_{i,t}) \cdot B_{i,t}}{(1+\rho_{i,t})^s \cdot \Pi_{t,s}} \right],$$

where $\Pi_{t,s} \equiv \prod_{j=1}^{s}(1+\pi_{t+j})$ is the gross cumulative inflation rate from loan origination at period t through s. The lender forms expectations over two types of risks. On one hand, there is idiosyncratic loan-level default risk: at each period s, the loan repays with probability $P_{i,t}$, in which case the lender earns the payment $D_{i,t,s}$, or it defaults with probability $1 - P_{i,t}$, in which case the lender earns the principal reduced by the loss given default, $(1 - LGD_{i,t})B_{i,t}$. On the other hand, the payments themselves are risky due to aggregate uncertainty, both due to inflation realized between origination and the period the payment is received $\Pi_{t,s}$ and, in the case of floating rate loans, to fluctuations in the underlying reference rate.

Assume now that the loan is a non-amortizing term loan, with each payment consisting of interest over the life of the loan, and the final payment consisting of a lump-sum principal repayment. Thus $D_{i,t,s} = r_{i,t,s} \cdot B_{i,t}$ for $s < T_{i,t}$ and $D_{i,t,T_{i,t}} = (1 + r_{i,t,T_{i,t}})B_{i,t}$. The interest rate $r_{i,t,s}$ is either a constant number, in the case of fixed-rate loans, or a fixed spread over a floating benchmark rate for floating rate loans. We can then approximate the lender's break-even condition at origination as:[7]

$$1 = \sum_{s=1}^{T_{i,t}} \left[\frac{P_{i,t}^s \cdot \mathbb{E}_t\left(r_{i,t,s}\right) + P_{i,t}^{s-1} \cdot (1 - P_{i,t}) \cdot (1 - LGD_{i,t})}{(1+\rho_{i,t})^s \cdot \mathbb{E}_t(\Pi_{t,s})} \right] + \frac{P_{i,t}^{T_{i,t}}}{(1+\rho_{i,t})^{T_{i,t}} \cdot \mathbb{E}_t(\Pi_{t,T_{i,t}})}, \quad (12)$$

This equation balances the present value of expected payments from the borrower against

[7] This is an approximation, as we abstract from the covariance between the expected floating rate and expected inflation, and from the Jensen's inequality term that arises because inflation appears in the denominator. Accounting for these terms would significantly complicate the estimation $\rho_{i,t}$, as it would require information about the expected term structure of the covariance between reference rates and inflation.

the lender's opportunity cost, ensuring that the lender breaks even. For a fixed-rate term loan, data on $(P_{i,t}, LGD_{i,t}, T_{i,t}, r_{i,t}, \{\mathbb{E}_t(\Pi_{t,s})\}_{s=1}^{T_{i,t}})$ allows us to solve this equation for the match-specific lender's discount rate $\rho_{i,t}$.

Floating Rate Loans. The data contains both fixed and floating rate loans. To estimate $\rho_{i,t}$ for floating rate loans, it is necessary to obtain estimates of $\mathbb{E}_t(r_{i,t,s})$, the expected interest rate. Floating rate loans typically charge a reference rate plus a spread. For our analysis, we use smoothed daily yield curve estimates provided by the Federal Reserve Board, based on the methodology described in Gürkaynak et al. (2007). Under the expectations hypothesis, long-term interest rates are assumed to reflect the market's expectations of future short-term rates. For each floating rate loan, we compute the sequence of forward short-term interest rates at the date of origination, and add the (fixed) loan spread to obtain a sequence of interest rates that are used to price that loan. Using this framework, we back out $\mathbb{E}_t(r_{i,t,s})$ for each loan by combining the treasury forward rate with the loan's spread.[8] It is worth noting that the majority of floating rate loans in our sample are indexed to the LIBOR/SOFR rather than Treasury rates. However, for the period in analysis, the spread between the SOFR and short-term Treasury rates is negligible. In the absence of readily available forward curve estimates for the LIBOR or SOFR, we treat them as identical to the Treasury curve.

Interpreting Loan Values and Loss Given Default. How should loan values in the data map to our model? It is straightforward to map between the model and data, but only at loan origination. A firm that sells debt for the first time receives $Q \cdot b'$ in exchange. Thus, loan value recorded at origination corresponds to $Q \cdot b'$ in the model. Another important variable from the data is the expected loss given default (LGD), which is given as a percentage. At origination, we interpret LGD as the expected losses relative to the loan's book value, so

[8]More specifically, the estimate for the reference rate n years ahead at time t is given by $f_t(n, 0) = \beta_0 + \beta_1 \exp(-n/\tau_1) + \beta_2(n/\tau_1)\exp(-n/\tau_1) + \beta_3(n/\tau_2)\exp(-n/\tau_2)$ (Equation 21 in Gürkaynak et al. (2007)), where estimates for $(\beta_1, \beta_2, \beta_3\tau_1, \tau_2)$ at each date are regularly updated by the Board of Governors of the Federal Reserve and published at `https://www.federalreserve.gov/data/nominal-yield-curve.htm`. We compute the sequence of forward rates at loan origination, and add the fixed spread to obtain an estimate for the interest rate at each repayment point in time.

that $1 - LGD = \phi(k')/(Q \cdot b')$.

After origination, banks do not update loan values on the books to reflect changes in market prices. Instead, they use historical-cost accounting, recording loans at their original value and adjusting over time based on the stated interest rate and repayment schedule (Begenau et al., 2026). As a result, observed loan values after origination no longer correspond to the market value of debt, $Q_t \cdot b_{t+1}$. For this reason, we focus only on newly originated loans.

Lender's Discount Rate. For each loan origination, the lender's discount rate $\rho_{i,t}$ can be estimated by numerically solving equation (12). This requires credit registry data on: i) loan maturity $T_{i,t}$, ii) repayment probability $P_{i,t}$, iii) loss given default $LGD_{i,t}$, and iv) interest rate or spread. For floating rate loans, data on spreads is combined with the Treasury forward curve to obtain $\mathbb{E}_t(r_{i,t,s})$. Finally, we compute $\{\mathbb{E}_t \Pi_{t,s}\}_{s=1}^{T_{i,t}}$ for each loan using estimates for the expected inflation term structure from Federal Reserve Bank of Cleveland (2025).

An instructive case is that of a hypothetical case of a fixed real interest rate loan; this serves as a useful approximation and is described in the following proposition:

Proposition 2 (Lender's discount rate for a fixed real rate)**.** *For a fixed real interest rate loan, the lender's discount rate can be written as:*

$$1 + \rho_{i,t} = P_{i,t} \left(1 + r_{i,t}\right) + \left(1 - P_{i,t}\right) \left(1 - LGD_{i,t}\right). \tag{13}$$

where $r_{i,t}$ is the fixed real interest rate on the loan. This expression reflects the lender's return, accounting for repayment in non-default states and recovery in default states. In this case, $\rho_{i,t}$ is independent of the loan's maturity $T_{i,t}$, which simplifies its calculation and interpretation.

Firm's Cost of Capital. We now turn to the firm's cost of capital. Our next result provides a formula to measure r^{firm} directly in the data.

Proposition 3 (Firm's Cost of Capital)**.** *We can measure the firm's cost of capital as:*

$$1 + r_{i,t}^{firm} = (1 + \rho_{i,t}) - \underbrace{(1 - P_{i,t})(1 - LGD_{i,t})}_{Expected\ Recoveries}.$$

In this expression, $(1-P_{i,t})(1-LGD_{i,t})$ represents expected recoveries, capturing the key difference between $\rho_{i,t}$ and $r_{i,t}^{firm}$: lenders benefit from expected recoveries, but borrowers get zero payoff in the default state, regardless of whether the lender recovers anything on the loan. Thus the expected cost of the loan for the firm is lower than the expected return for the lender. This formula allows us to measure the firm's cost of capital for each loan in the data, at origination.

Social Cost of Capital. Finally, we explain how we use data to estimate the social cost of capital, $r_{i,t}^{social}$. For measurement, we specialize and assume that the liquidation technology is linear, meaning it takes the form $\phi_i(k_i) = \phi_i \cdot k_i$. Combining with Equation (11) and proposition 3, this yields a formula for $r_{i,t}^{social}$ in terms of objects in the data.

Proposition 4. *Assume a linear liquidation technology. The social cost of capital is then:*

$$\begin{aligned} 1 + r_{i,t}^{social} &= \left(1 + r_{i,t}^{firm}\right)\mathcal{M}_{i,t} + (1 - P_{i,t}) \cdot (1 - LGD_{i,t}) \cdot lev_{i,t} \\ &= (1 + \rho_{i,t})\,\mathcal{M}_{i,t} + (lev_{i,t} - \mathcal{M}_{i,t}) \cdot (1 - P_{i,t}) \cdot (1 - LGD_{i,t}) \end{aligned}$$

where $lev_{i,t} := \frac{Q_{i,t} \cdot b_{i,t+1}}{k_{i,t+1}}$ *is the firm's leverage ratio.*

In our empirical analysis, we set the price feedback multiplier $\mathcal{M}_{i,t} = 1$.[9] Under that calibration, the social cost of capital simplifies further:

$$1 + r_{i,t}^{social} = \underbrace{1 + \rho_{i,t}}_{\text{Lender Discount Rate}} + \underbrace{(\text{lev}_{i,t} - 1) \cdot (1 - P_{i,t}) \cdot (1 - LGD_{i,t})}_{\text{Financial Frictions}} \tag{14}$$

The social cost of capital is thus equal to the lender's discount rate, plus a term that reflects financial frictions related to default and recovery. This latter term reflects the tension between lenders and borrowers. The presence of this term also implies that whether the

[9] We later show that our results are robust to relaxing this assumption in Section 6.1.

social return on capital exceeds the lender's discount rate or not is a function of whether firm leverage exceeds 1 or not. This is due to the following: lenders care about average recovery per dollar of debt, $\phi_i(k_i)/(Q_i b_i)$, which is equal to $1 - LGD_i$ in the data. The planner, on the other hand, cares about the marginal recovery $\phi_i'(k_i)$, which is equal to $(1 - LGD_i) \cdot \text{lev}_i$ in the data. The two coincide when $\text{lev}_i = 1$, and so the social cost of capital equals the lender's discount rate in that case. In general, for firms with $\text{lev}_{i,t} \in (0, 1)$, we obtain the following ranking: $r_{i,t}^{firm} < r_{i,t}^{social} < \rho_{i,t}$.

5 Empirical Results

We now describe the results of applying the measurement exercise described in the previous section to the U.S. Y-14 data.

5.1 Summary Statistics

We provide summary statistics for key variables in Table 1. Since the unit of observation is a loan origination, all reported firm financials correspond to the financials reported in the quarter in which that origination took place. The average annual loan interest rate in our sample is 4.35%. Adjusted for expected inflation at origination, this results in an average real rate of 2.47%. We compute the real interest rate at origination by subtracting (annualized) expected inflation over the maturity of the loan, using the expected inflation term structure that prevailed at time of origination. These loans have an average expected default probability of 1.41% over the next year, and banks expect to lose, on average, 34.35% of the outstanding value of the loan in the event of default. As a result, the lender's discount rate, ρ, averages 1.96%.[10] The social cost of capital and firm's cost of capital respect the aforementioned ranking and are even lower, at 1.75% and 1.04% respectively.

Real interest rates vary across loans, with a standard deviation of 1.3% (1.8% for nominal

[10] The negative covariance of r and P means that the average ρ is lower than we might have expected from the raw averages of P, r, and LGD. To see this, consider the approximation for real fixed rate loans, $1 + \rho = P(1 + r) + (1 - P)(1 - LGD)$; the average value of $P(1 + r)$ will be brought down by the fact that P is low when $(1 + r)$ is high.

rates), reflecting heterogeneity both within and across time. The lender's discount rate shows similar heterogeneity, with a standard deviation of 1.5%. In contrast, the social cost of capital and firm's cost of capital have higher heterogeneity, with standard deviations of 1.7% and 2.6%, respectively.

Why are the variances of the lender's discount rate and the contractual interest rate similar? To build intuition, we can focus on the formula for the lender's discount rate, ρ, for loans with a fixed real rate in Equation (13). With some rearrangement, the lender's discount rate can be expressed as $\rho = r - (1 - P)(r + LGD)$. Since r is small at annual frequencies compared to LGD, we can further approximate $\rho \approx r - (1 - P) \cdot LGD$. This yields the following variance decomposition:

$$\mathrm{Var}\,(\rho) \approx \mathrm{Var}\,(r) + \mathrm{Var}\,((1 - P) \cdot LGD) - 2 \cdot \mathrm{Cov}\,(r, (1 - P) \cdot LGD) \tag{15}$$

The variance of ρ is similar to the variance of r because the variance of expected losses,$(1 - P) \cdot LGD$, is offset by the covariance term: interest rates are higher when the lender's expected losses are high.

We view these results as a vindication for our method of estimating the cost of capital. If our observed measures of default probabilities and recovery rates were just noise, then the variance of the lender's discount rate would be substantially greater than the variance of interest rates: the covariance term would be zero, and the term $\mathrm{Var}\,((1 - P) \cdot LGD)$ would push the variance of ρ substantially above the variance of r. Instead, the variance of the cost of capital is similar to that of the interest rate, suggesting that default probabilities and recovery rates covary with interest rates in the way that we would expect.

5.2 Interest Rates, Default, and Recovery Rates

We begin by showing the raw data that goes into our main formulas: interest rates, default probabilities, and expected loss given default. We first show the time series by quarter of origination of different percentiles of the distribution in Figure 1. The first panel plots real rates at origination for all loans, while the second panel plots contractual interest

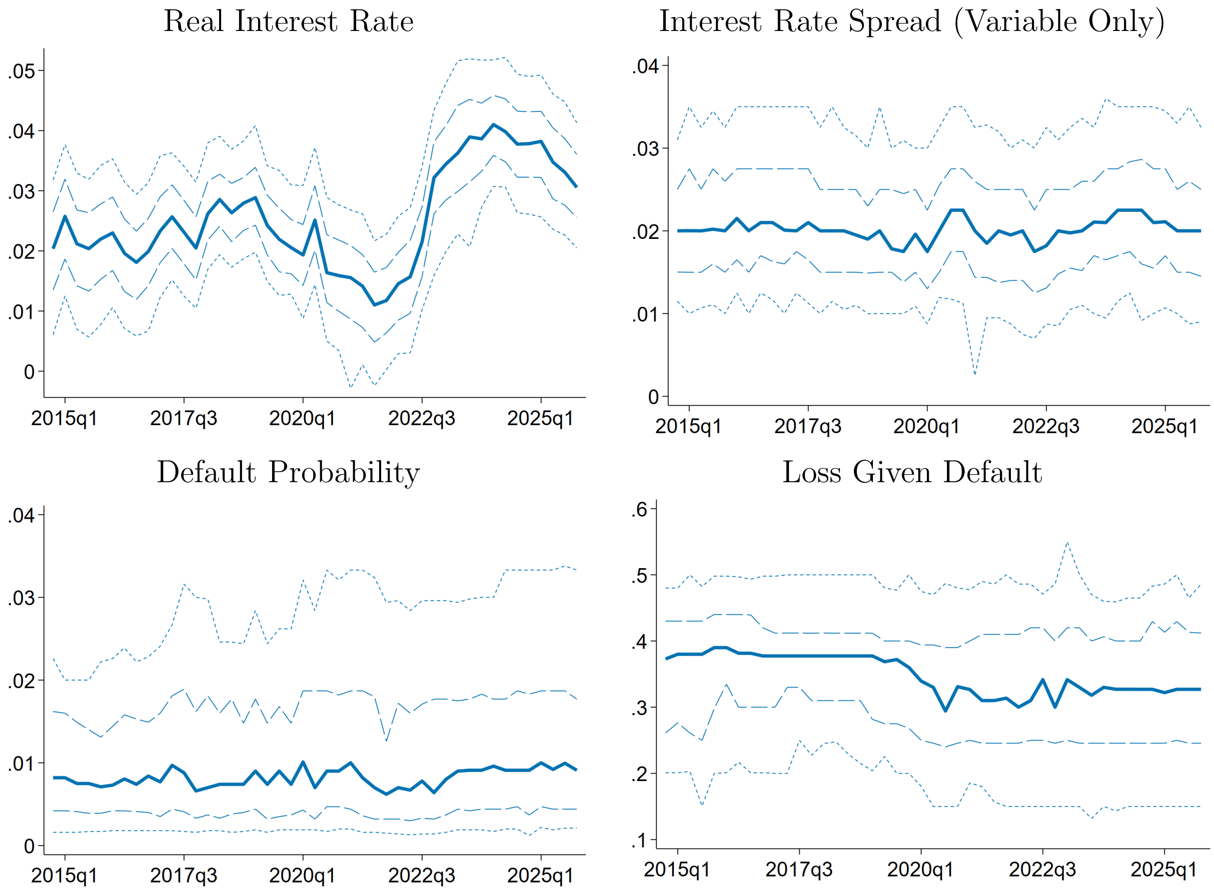
Table 1: Summary Statistics

	Mean	SD	p10	p50	p90
Interest Rate (%)	4.35	1.77	2.22	4.07	6.80
Maturity (Years)	6.81	4.62	3.00	5.00	10.25
Real Interest Rate (%)	2.47	1.27	0.91	2.42	4.09
ρ (%)	1.96	1.53	0.45	1.96	3.74
r^{firm} (%)	1.04	2.59	-0.80	1.34	3.15
r^{social} (%)	1.75	1.72	0.15	1.81	3.59
Probability of Default (%)	1.41	2.32	0.17	0.82	2.94
Loss Given Default (%)	34.35	12.97	16.20	36.00	49.80
Loan Amount ($ Millions)	11.68	70.48	1.10	2.56	25.00
Sales ($ Millions)	1,396.68	6,628.60	2.38	62.83	1,791.58
Assets ($ Millions)	2,036.80	9,893.13	1.02	38.63	2,372.00
Leverage (%)	71.42	24.54	42.02	70.49	100.00
Return on Assets (%)	27.28	54.94	4.43	15.59	48.50
N Loans	67089				
N Firms	39141				
N Fixed-Rate Loans	33927				
N Variable-Rate Loans	33162				

Notes: Data are for our full sample, and each observation is a newly originated loan. Loan maturity is reported in years. Loan amount, sales, and assets are reported in millions of dollars. All other variables are reported in percent. Firm financial variables are measured in the quarter of origination. The real interest rate is the contractual interest rate minus annualized expected inflation, computed over the maturity of the loan. Probability of default is the bank's estimate of the one-year probability of default. Loss given default is the bank's expected losses on the loan in the event of default, measured as a percentage of the loan's value. The columns report the mean, standard deviation, and 10th, 50th, and 90th percentiles.

rate spreads for floating-rate loans only. During the time period we study, interest rates fall and then rise concurrent with the movements of monetary policy, while average spreads are relatively stable. Default probabilities show a modest upward secular trend, along with a temporary spike around the time of the COVID-19 pandemic. Expected losses given default fall around the onset of the pandemic, implying that banks expect larger recoveries in the event of default. Note, however, that the magnitude of the change in recoveries is sufficiently small that it has little effect on ρ, since this change is multiplied by the (small) probability of default.

Figure 1: Interest Rates and Credit Risk

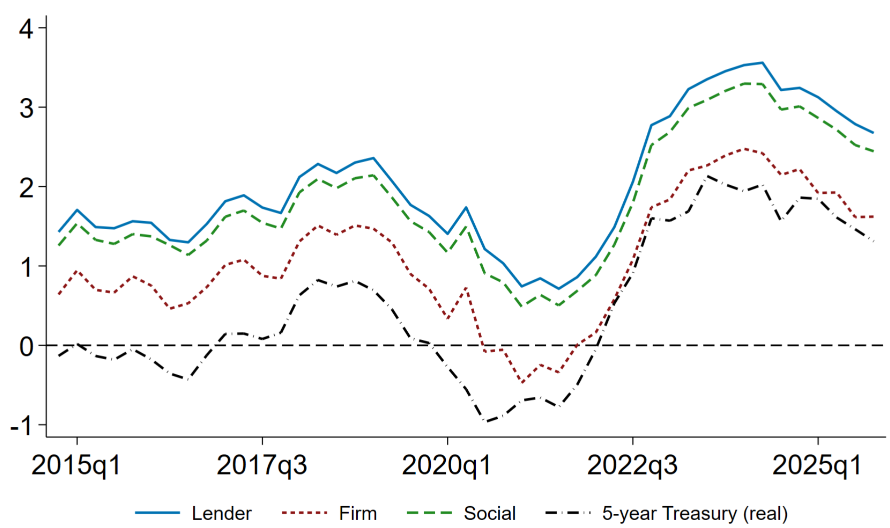


Notes: Each panel plots the distribution of the indicated variable by quarter of loan origination: real interest rate at origination, interest rate spread at origination, probability of default, and loss given default. The interest-rate-spread panel includes only floating-rate loans. The solid line reports the median, while the dashed lines report the 10th, 25th, 75th, and 90th percentiles.

5.3 Cost of Capital

Next, in Figure 2, we plot the lender's discount rate, ρ, the firm's cost of capital, r^{firm}, and the social cost of capital, r^{social}, against the real five-year Treasury rate. All rates covary strongly with the real five-year Treasury rate. The average lender's discount rate is similar to the average r^{social}, with a roughly 20 basis points spread. The lender's discount rate remains above the real five-year Treasury rate, although it has a delayed reaction to movements in Treasury rates: the spread rises when Treasury rates fall and falls once Treasury rates rise again. Note that the lender's discount rate is already adjusted for default risk, and so this cannot explain the spread relative to treasuries. While the social cost of capital, r^{social} is quite close to ρ, the firm's cost of capital, r^{firm}, tracks the Treasury rate more closely with a small spread.

Figure 2: Cost of Capital and Treasury Rate



Notes: The figure plots averages by quarter of origination for the lender's discount rate (ρ), firm cost of capital (r^{firm}), and social cost of capital (r^{social}). The real five-year Treasury rate is plotted for comparison. All rates are annual and reported in percent.

While our analysis takes into account the maturity of the loan, there are potential concerns that loans of different maturities may face different rates, even if they reflect a constant spread on a (time-varying) risk-free rate. To mitigate these concerns, in Appendix B.5, we redo our analysis focusing on loans with a 5-year maturity. This is the most common maturity for fixed rate loans. Five-year loans are convenient because they allow direct comparison to the five-year Treasury rate. The average cost of capital for five-year loans is very similar to that of the overall sample.

5.4 Cross-Sectional Heterogeneity

While the means display similar movements, they mask a substantial amount of cross-sectional heterogeneity in these measures. In this section, we show that this heterogeneity is substantial across measures even after controlling for time, lender and firm fixed effects. To this end, we follow the variance decomposition of Daruich and Kozlowski (2023). To ensure that we can estimate firm level fixed effects we subset our sample to the set of firms with five or more distinct loans. We then progressively add time, lender-time, and firm-lender-time fixed effects, building to the fixed-effects specification displayed in Equation (16) below, where i indexes the firm, τ represents the quarter of origination, b indexes the lender (BHC), and l represents the particular loan.

$$r_{\tau bil} = \alpha_\tau + \gamma_{\tau b} + \delta_{\tau bi} + \varepsilon_{\tau bil} \tag{16}$$

The results are in Table 2, for the nominal interest rate, real rate, lender's discount rate, firm cost of capital, and social cost of capital. The time fixed effect explains 72% of the variance in nominal interest rates, 52% of the variance in real interest rates, and 48% of the variance in the lender's discount rate. For all five variables, adding in lender-time fixed effects explains a negligible share of the variance (between 2 and 5%), suggesting that heterogeneity across lenders is not an important source of heterogeneity in interest rates or the different measures of cost of capital. Adding in firm-lender-time fixed effects explains an additional 13% to 29% of the variance of these measures. Finally, notice that the loan-level variance, after controlling for firm-lender-time fixed effects, remains substantial, ranging from 13% for nominal interest rates to up to 48% for the firm cost of capital. To the extent that our measure of misallocation depends on the variance of r^{social}, these results show that a significant share of that variance is loan-specific and cannot be accounted for by either time, lender, or borrower effects.

Table 2: Variance Decomposition

	Time	Bank	Firm	Loan
Contractual Rate	72.06	1.62	13.01	13.31
Real Rate	52.00	3.78	23.26	20.96
ρ	48.42	4.01	20.58	26.98
r^{firm}	19.04	4.07	28.82	48.07
r^{social}	40.35	4.61	22.41	32.64

Notes: Each entry reports the additional percentage of the variable's total variance explained by adding in fixed effects of that type. Following Equation (16), Time is the share explained by quarter-of-origination fixed effects, Bank is the additional share explained by including bank-quarter fixed effects, Firm is the additional share explained by including firm-bank-quarter fixed effects, and Loan is the residual variance share. The sample is restricted to firms with at least five distinct loans in our data, and contains 1.982 firms and 17,622 loans. Entries in each row sum to 100, up to rounding.

Appendix B.3 further explores the correlation between the cost of capital and firm-level covariates such as leverage, return on assets, and assets. While the effects of leverage and size vary across the specific measure, there is a consistent positive correlation between all measures of cost of capital and the return on assets: the cost of capital is consistently higher at firms with high return on assets, even though the return on assets is negatively correlated with the real interest rate once leverage and size are controlled for. Although we cannot attach a causal interpretation to the estimated coefficients, this would be consistent with a model where causality runs from the cost of capital to firm decisions: firms with a higher cost of capital will demand a high return on their investments.

5.5 Misallocation

What does the heterogeneity in the cost of capital imply for the cost of misallocation? To answer this question, we use our formula for misallocation from Corollary 1. Applying this formula requires calibrating two parameters: $\mathcal{E}$ and δ. In a Cobb-Douglas setting, with $f(k,z) = z \cdot k^{\alpha}$ and no default, the elasticity simplifies to $\mathcal{E} = \frac{\alpha}{1-\alpha}$. We set $\mathcal{E} = \frac{1}{2}$, which is consistent with $\alpha = \frac{1}{3}$. We set the annual depreciation rate to $\delta = 0.06$.[11] We compute

[11] Note that in steady state $\delta = \frac{I/Y}{K/Y}$. In the data, at the annual frequency, the capital-output ratio is about 3 while the investment-output ratio is about 0.18 (we measure capital as BEA Current-Cost Net Stock

this statistic by quarter of origination, in order to focus on within-period misallocation. Our model does not contain aggregate shocks, and we would thus need a richer model to study misallocation across time. We interpret our results below as reflecting what misallocation would be for an economy that remained in the same steady state.

We plot our estimates in Figure 3. We find that, overall, the misallocation of capital resulting from heterogeneity in the cost of capital, plus the financial friction term, is small. In the period before the COVID pandemic the implied misallocation is low and decreasing: reallocating capital across firms would increase aggregate output by 0.71%, on average, during this period. This number rises with the onset of the pandemic, averaging 1.62% during 2020 and 2021, before falling back to 1.16% starting in 2022.

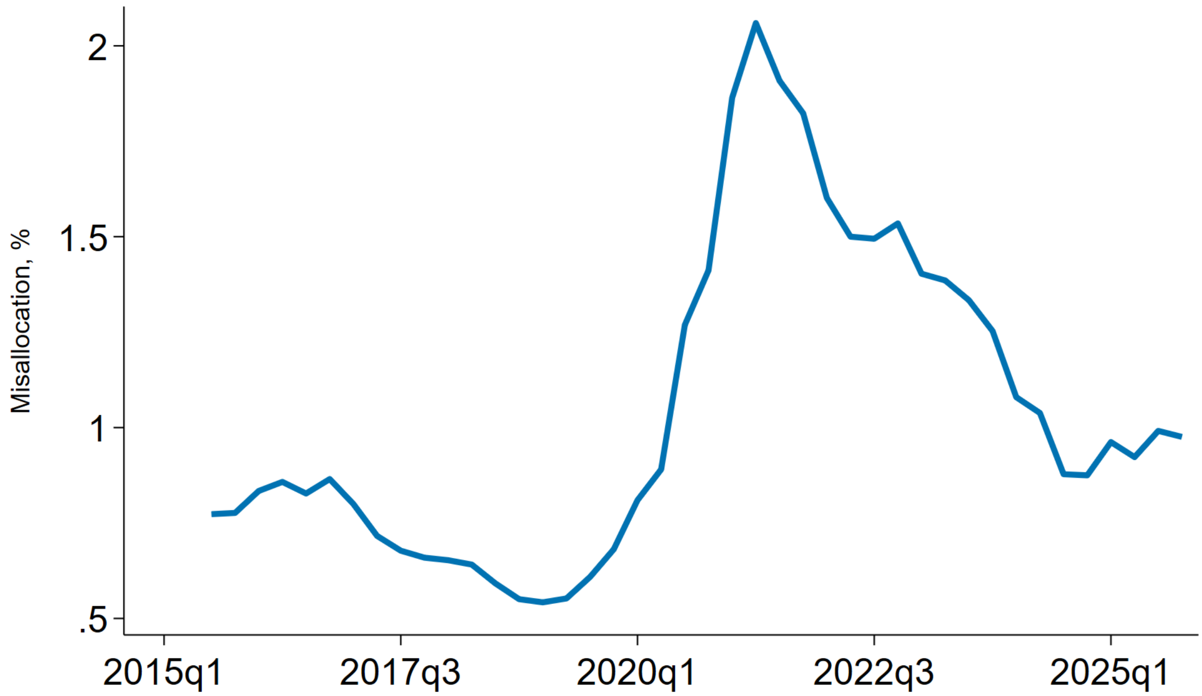
5.6 Decomposing Misallocation

To understand the drivers of misallocation, we use Equation (14) to decompose our measure into the component coming from heterogeneity in lender discount rates, ρ, and the component coming from heterogeneity in the financial frictions term. We perform two counterfactuals. In the first, we replace ρ with its average value for that quarter, which is informative about how much misallocation arises from heterogeneity in the financial frictions term. In the second counterfactual we set the financial frictions term equal to its average value within the quarter, allowing us to measure the misallocation that arises from heterogeneity in ρ. Note that neither counterfactual changes the within-quarter average of r^{social}, and thus the results are exclusively driven by changes in the variance of r^{social}.

We show the results of these decompositions in Figure 4, for three distinct time periods. The main takeaway is that misallocation is mostly driven by heterogeneity in lender discount rates. In the pre-pandemic period, if ρ is equalized across firms then the cost of misallocation falls to just 0.08% instead of 0.71%. In contrast, the cost of misallocation in the counterfactual with a constant financial friction is 0.54%. Note that there is an interaction

of Fixed Assets and investment is GDPI in FRED). Hence, at an annual frequency, $\delta = 0.06$. Additionally, recall that, as previously explained, $\mathcal{M} = 1$.

Figure 3: Misallocation over Time

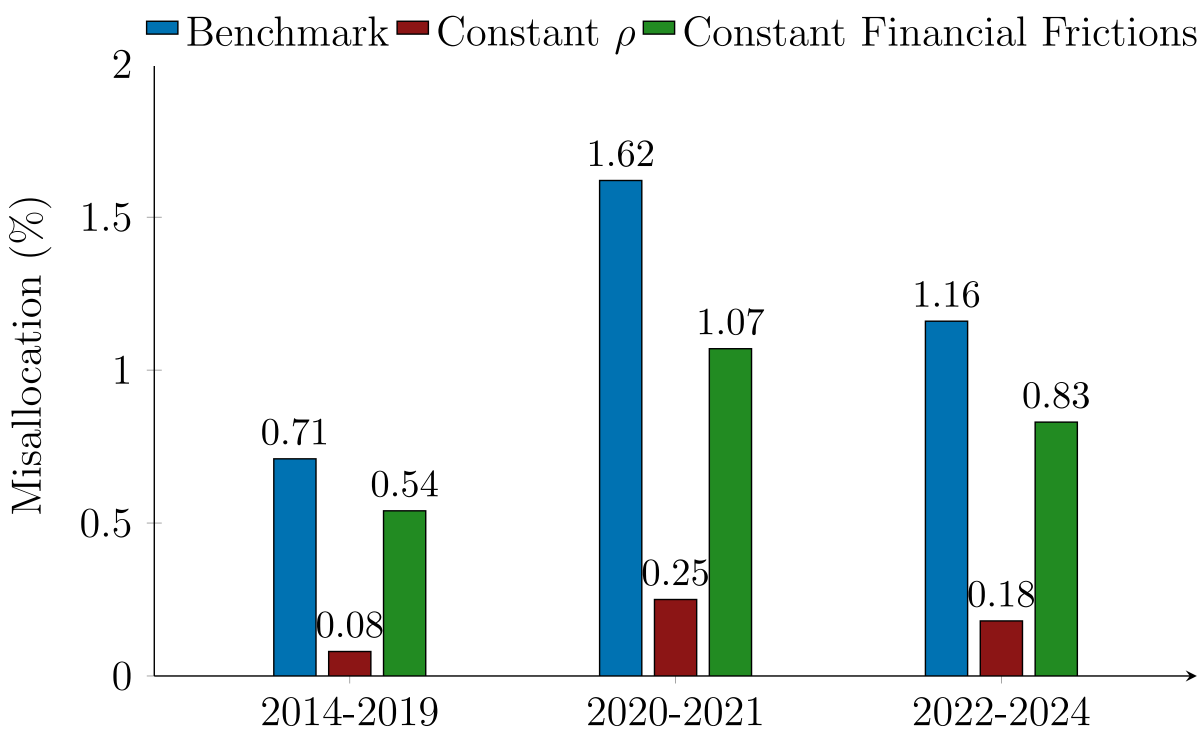


Notes: The figure plots the implied cost of misallocation by quarter of origination, using the formula in Corollary 1. The cost of misallocation is the percentage increase in output in a counterfactual economy where the planner is free to reallocate physical capital across firms, keeping exit decisions fixed. The series is reported as a four-quarter moving average.

between ρ and the financial frictions term as total misallocation exceeds the sum of the two counterfactuals. During the pandemic period (2020-2021), there is a significant increase in total misallocation, which more than doubles. While still small, the financial frictions term increases significantly from 0.08% to 0.25%. Quantitatively, the large increase comes from an increased dispersion in ρ. In the post-pandemic period, misallocation remains over 60% more elevated than in the pre-pandemic period. Once again, ρ plays the dominant role: misallocation would be only 0.18% in the constant ρ counterfactual.

Appendix B.5 repeats our misallocation analysis focusing on loans with a maturity of five years. We find very similar results to those of the main analysis, which we take as evidence of their robustness, as it confirms that heterogeneity in the cost of capital across firms is not

Figure 4: Decomposing the Cost of Misallocation



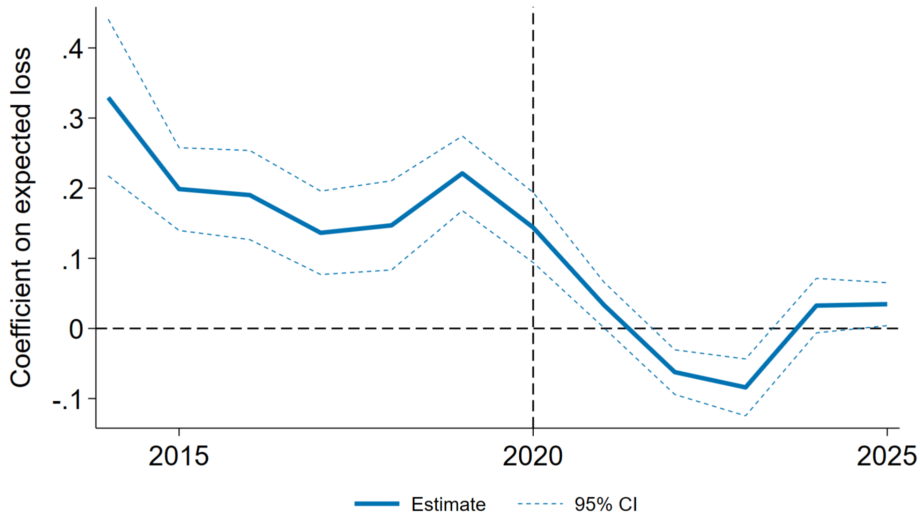
Notes: The "Benchmark" bars report the average cost of misallocation from Figure 3 within each period. The "Constant ρ" bars report the average cost of misallocation in a counterfactual exercise where we set the estimated ρ for each newly originated loan equal to the average ρ for that quarter. The "Constant Financial Frictions" bars report the average cost of misallocation in a counterfactual exercise where we set the second term in Equation (14) equal to its average value for that quarter.

driven by differences in maturity or term structure.

5.7 The 2020–2021 Increase in Misallocation

Our analysis in Figure 4 reveals that while both components of the social cost of capital contribute to the increase in misallocation in 2020–2021, the heterogeneity in ρ is quantitatively the most important factor. We can further decompose the heterogeneity in ρ using the approximation we derived earlier, $\rho \approx r - (1-P) \cdot LGD$. Empirically, r^{social} and ρ are very close, so $\text{Var}(\rho) \approx \text{Var}\left(r^{social}\right)$. In an efficient benchmark, there is no misallocation of capital when $\text{Var}\left(r^{social}\right) = 0$. These two approximations imply that the interest rate r should increase one-for-one with expected losses $(1-P) \cdot LGD$ in an efficient economy.

Figure 5: Pass-Through of Expected Losses to Real Interest Rates



Notes: The solid line plots the coefficient estimates from regressions of real interest rates on expected losses, as described in Equation (17). Expected losses are defined as the probability of default multiplied by the expected loss given default. The regressions are estimated separately by year of origination and control for quarter fixed effects. Dashed lines plot 95% confidence intervals.

To examine the relationship between interest rates and expected losses, we estimate the following regression separately for each year y:

$$r_{it} = \alpha_t + \beta_y \cdot \text{Expected Losses}_{it} + \varepsilon_{it} \tag{17}$$

where i indexes loans, t indexes the quarter of origination, the dependent variable r_{it} is the real interest rate at origination, and α_t are quarter-of-origination fixed effects. Expected losses are measured as $(1 - P_{it}) \cdot LGD_{it}$ at origination. For each year y, we pool all newly originated loans made in that calendar year to estimate β_y. This pooling reduces estimation noise relative to estimating a separate β for each quarter. We nevertheless include quarter-of-origination fixed effects to account for within-year changes in the average interest rate.

We plot the estimated regression coefficients, by year, in Figure 5. The results are striking in two respects. First, although there is a clear statistically significant relationship between expected losses and the real interest rate before the COVID period, the relationship is much weaker than the efficient benchmark would predict. Whereas we would expect $\beta \approx 1$ under the efficient benchmark, we instead see $\beta \approx 0.2$ in the pre-COVID period. Second, starting in 2020, there is a steep decline in β. By 2021, $\beta \approx 0$, and β is negative and statistically significant in 2022 and 2023.[12]

In light of the results in Figure 5, we focus on three factors that can generate dispersion in the cost of capital. First, the weak relationship between expected losses and interest rates generates dispersion in the cost of capital. Second, since interest rates do not increase one-for-one with expected losses, an increase in the variance of expected losses increases the dispersion of the cost of capital. Third, the weakening of the relationship between expected losses and interest rates also increases the dispersion of the cost of capital.

To better understand how these forces drive the post-COVID rise in misallocation, we run the following regression separately for each quarter among newly originated loans:

$$r_{it}^{social} = \alpha + \beta \cdot \text{Expected Losses}_{it} + \varepsilon_{it} \tag{18}$$

We use this regression to decompose the post-COVID rise in misallocation of credit into four sources: changes in (α, β), changes in the variance of expected losses, changes in the mean of expected losses, and changes in the variance of the residual ε_{it}. To measure the contribution of each source to the rise in misallocation, we use Equation (18) to compute counterfactual values of r_{it}^{social} under different alternative scenarios, and then recompute misallocation. We compute (α, β), $\mathbb{E}\left[\text{Expected Losses}_{it}\right]$, $\text{Var}\left(\text{Expected Losses}_{it}\right)$, and $\text{Var}\left(\varepsilon_{it}\right)$ in each quarter, take the average of each of those values for the period, and then compute misallocation as:

[12] It would be difficult for measurement error to explain these patterns. If expected losses are measured with noise, then this will attenuate our estimate of β towards zero. However, explaining the drop of β to zero post-COVID would require that the measure of expected losses suddenly became almost entirely noise post-COVID. Moreover, since measurement error attenuates estimates toward zero, it cannot explain the negative estimates in 2022 and 2023.

$$
\begin{aligned}
\frac{Y^*}{Y} &= \exp\left(\frac{1}{2}\cdot\mathcal{E}\cdot\log\left(1+\frac{\mathrm{Var}\left(r^{social}\right)}{\mathbb{E}\left[r^{social}+\delta\right]^2}\right)\right) \\
&= \exp\left(\frac{1}{2}\cdot\mathcal{E}\cdot\log\left(1+\frac{\beta^2\cdot\mathrm{Var}\left(\text{Expected Losses}_{it}\right)+\mathrm{Var}\left(\varepsilon_{it}\right)}{\left(\alpha+\beta\cdot\mathbb{E}\left[\text{Expected Losses}_{it}\right]+\delta\right)^2}\right)\right)
\end{aligned}
$$

Since misallocation is a nonlinear function of the sources we are interested in, we use the Shapley-Owen-Shorrocks decomposition (Shorrocks, 2012). This decomposition begins with the 2014–2019 values of each component, and then replaces them sequentially with the 2020–2021 values, averaging across all possible orders.

We find that the relationship between expected losses and the cost of capital is essential for understanding the rise in misallocation. Specifically, we find that the rise in $\mathrm{Var}\left(\text{Expected Losses}_{it}\right)$ explains 41% of the rise in misallocation. Because interest rates only rise modestly with expected losses, there is a strong negative relationship between expected losses and r^{social}, and so a rise in the variance of expected losses leads to higher misallocation. We also find that the weakening of the relationship between interest rates and expected losses is important: the change in (α, β) explains 42% of the rise in misallocation. This is partly because a weaker relationship between interest rates and expected losses implies a higher variance of r^{social}, and partly because changes in α, β lead to a lower $\mathbb{E}\left[r^{social}\right]$ in the denominator. Finally, the remaining factors are much less important: an increase in $\mathrm{Var}\left(\varepsilon_{it}\right)$ explains 14% of the rise in misallocation, and changes in $\mathbb{E}\left[\text{Expected Losses}_{it}\right]$ explain just 3%.

In summary, the post-COVID spike in misallocation was driven by two closely related forces: an increase in the dispersion of expected losses and a weakening of the extent to which interest rates reflected those losses. Put differently, interest rates became less informative signals of underlying credit risk precisely when risk heterogeneity increased. We do not take a stance on the mechanisms behind this weakening relationship. Possible explanations include expectations of government support through programs similar to the Paycheck Protection Program (PPP), stronger incentives to evergreen loans by extending credit at below-market rates (Faria-e-Castro et al., 2024), or other changes in credit-market conditions during this period.

6 Extensions and Robustness

In this section, we present a variety of robustness checks and extensions. First, we relax the assumption that the price-impact term is equal to $\mathcal{M} = 1$: we estimate it directly using security- and firm-level financial data, and then recompute our measure of misallocation allowing for variation in this term. Second, we recompute our benchmark measure using alternative weights for r^{social} that reflect the size of each loan. Third, we adjust r^{social} for estimated exposures to aggregate risk. We show that the magnitude and dynamics of our baseline measure are robust to any of these modifications. Fourth, we validate our measures of cost of capital by showing that they are related both to the average revenue product of capital, as in Hsieh and Klenow (2009), and to text-based measures of the cost of capital, as in Gormsen and Huber (2023). Finally, we perform a cross-country comparison of misallocation using an approximation of our measure, applied to summary statistics reported in other papers using credit-registry data from Pakistan, Brazil, and Mexico.

6.1 Robustness

Heterogeneous $\mathcal{M}$. As explained in Section 4, we set the price feedback multiplier to $\mathcal{M} = 1$ for our baseline measure. In Appendix B.2, we relax this assumption and estimate a loan-specific price feedback multiplier, $\mathcal{M}_i$. This procedure yields a distribution of $\mathcal{M}_i$ that is extremely concentrated around 1, which supports our baseline assumption. Nevertheless, we recompute r_i^{social} and the measure of misallocation using this distribution of price feedback multipliers, and again find that the magnitude and dynamics of misallocation are similar to those in the baseline.

Weighting. Our baseline measure is unweighted, which Hughes and Majerovitz (2025) argue is a good approximation. Appendix B.4 replicates our measure while weighting the mean and variance of r^{social} by the size of the corresponding credit facility. The resulting measure of misallocation is very similar to the baseline measure, both in levels and in dynamics over the sample.

Allowing for Labor Reallocation. The only input in our model is capital, and thus we implicitly hold the allocation of labor fixed when we measure the cost of misallocation. We can adjust our measure of misallocation to allow for labor reallocation. Suppose that, as in Hsieh and Klenow (2009), firms have the Cobb-Douglas production function $f(k_i, l_i, z_i) = z_i k_i^{\alpha} l_i^{1-\alpha}$. Firms' output is aggregated CES, with elasticity of substition σ across varieties. Aggregate labor and aggregate capital are each held fixed. We will assume that the firm's choice of labor is undistorted: the only distortion comes from heterogeneity in r^{social}. If productivity and wedges are jointly log-normal, then we can apply the results of Hsieh and Klenow (2009) to show:

$$\log Y^* - \log Y = \frac{1}{2} \cdot \underbrace{\left(\sigma\alpha^2 + \alpha(1-\alpha)\right)}_{\mathcal{E}^{\text{new}}} \cdot \log\left(1 + \frac{\text{Var}\left(r_{i,t-1}^{social}\right)}{\mathbb{E}\left[r_{i,t-1}^{social} + \delta\right]^2}\right) \tag{19}$$

This formula is the same as our main formula, but our original $\mathcal{E}$ has been replaced with $\mathcal{E}^{\text{new}} = \sigma\alpha^2 + \alpha(1-\alpha)$. Thus, any of our misallocation numbers can simply be rescaled to account for labor misallocation.

In our main analysis, we calibrated $\mathcal{E} = \frac{1}{2}$ to match $\alpha = \frac{1}{3}$. Because the log output loss is proportional to $\mathcal{E}$, the estimated losses can be rescaled by $\mathcal{E}^{\text{new}}/\mathcal{E}$. If we use an elasticity of substitution $\sigma = 3$, as in Hsieh and Klenow (2009), then this yields $\mathcal{E}^{\text{new}} = \frac{5}{9}$, which is very close to our original $\mathcal{E}$. Using a very high elasticity of substitution, $\sigma = 10$, we get $\mathcal{E}^{\text{new}} = \frac{4}{3}$. This would increase losses of 1% to 2.7%. Thus, allowing for labor reallocation will increase the misallocation that we measure, but will not affect the substantive conclusion that the cost of misallocation is small in our setting.

6.2 Risk Adjustment

A potential concern is that dispersion in the lender discount rate, ρ, may partly reflect compensation for exposure to aggregate risk rather than misallocation. This concern is closely related to the work of David et al. (2022), who show that risk premia can generate dispersion in marginal products even in efficient allocations. To evaluate this possibility, we implement a risk-adjustment exercise that extends the two-step logic of Fama-MacBeth regressions (Fama and MacBeth, 1973) to our setting. In the standard approach, asset returns are first regressed on aggregate factors to estimate asset-specific factor loadings, and then cross-sectional returns are regressed on those loadings to estimate factor prices.

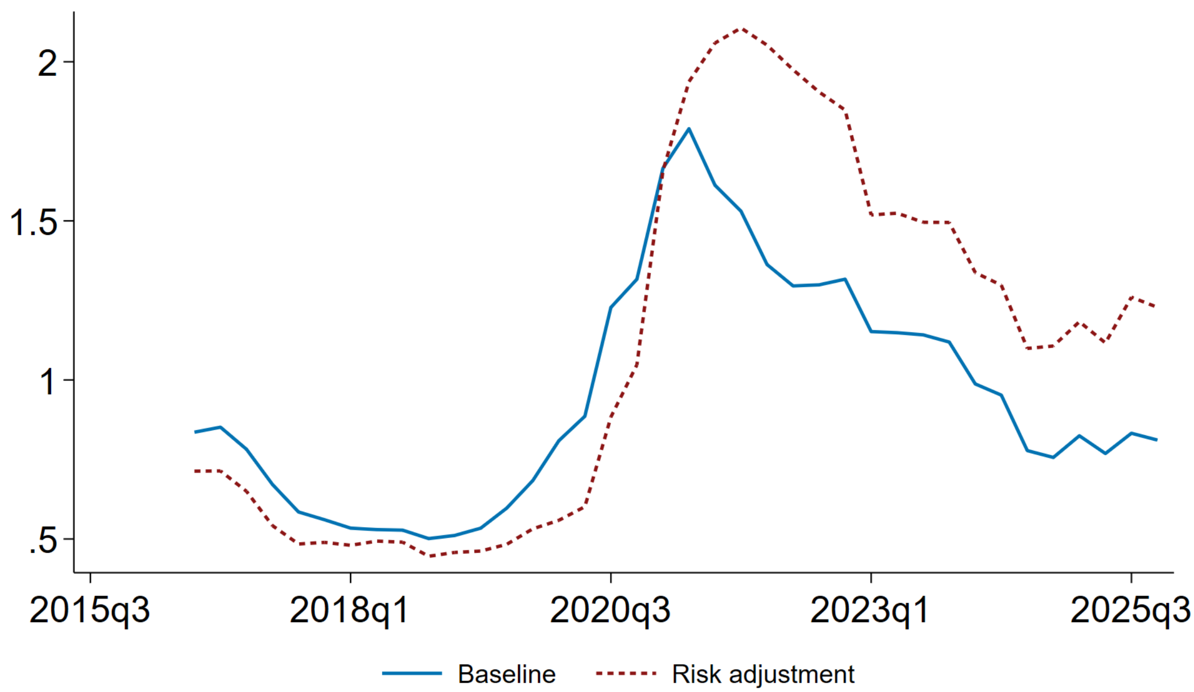
Our implementation has additional steps because Y-14 loans are not traded assets. We first estimate Fama-French three-factor loadings for corporate bonds, using duration-adjusted bond returns and the market, value, and size factors. Separately, we estimate stock betas and use monthly Fama-MacBeth regressions in the cross section of stocks to estimate the prices of these aggregate risk factors. We use stock returns for this step because the traded-equity sample is much larger and more liquid than the traded-bond sample. Following the risk-adjustment approach in Chiavari et al. (2025), we smooth the estimated factor prices with a three-year moving average.

We then map the stock-based betas into predicted bond betas for Compustat firms, merge these predicted bond betas to the Y-14, and project them on firm characteristics observed in the Y-14, including size, sales, cash, leverage, profitability, and industry. This projection lets us predict factor loadings for all Y-14 borrowers. Appendix C reports the exact regressions used in this procedure. Finally, for each Y-14 loan we subtract the annualized predicted aggregate-risk compensation from the measured lender discount rate,

$$\rho_{i,t}^{RA} = \rho_{i,t} - \left[\left(1 + \widehat{\beta}_{i,t}' \overline{\gamma}_t \right)^{12} - 1 \right],$$

and recompute the firm and social costs of capital, and the implied misallocation measure, using $\rho_{i,t}^{RA}$ in place of $\rho_{i,t}$.

Figure 6: Risk-Adjusted Misallocation



Notes: The figure compares the baseline measure of misallocation to the risk-adjusted measure. The risk-adjusted measure is constructed using three-year moving averages of the estimated factor prices. The cost of misallocation is reported as the percentage increase in output from efficiently reallocating capital across firms. Both misallocation series are reported as four-quarter moving averages. The baseline measure differs slightly from that in Figure 3 due to sample differences, as the risk-adjusted measure is not available for all loans.

Figure 6 compares the risk-adjusted measure to the benchmark. The two series are close overall, but the adjustment has different implications before and after COVID-19. Before COVID-19, the risk-adjusted measure is slightly below the benchmark, suggesting that part of the pre-pandemic dispersion in lender discount rates reflects compensation for aggregate risk, as emphasized by David et al. (2022). Starting with COVID-19, however, the risk-adjusted measure lies above the benchmark. This means that subtracting the estimated compensation for aggregate risk raises the measured pandemic increase rather than attenuating it. Aggregate-risk compensation therefore cannot explain the COVID-era rise in misallocation. This pattern suggests that the COVID-era rise reflects dispersion in lending terms beyond standard compensation for aggregate risk exposure, consistent with mispricing of risk or a deterioration in the allocation of credit.

6.3 Relation to ARPK-based Measures

The dominant approach to measuring misallocation builds on Hsieh and Klenow (2009) and typically consists of using data on firm financials to approximate the marginal revenue product of capital (MRPK) at the firm-level. Under certain assumptions about the structure of the model, researchers can approximate MRPK with the average revenue product of capital (ARPK). Aggregate measures of misallocation are then compute based on moments of the distribution for ARPK. Our approach, instead, uses data on the cost of borrowing to measure r^{social}, which we argue is informative about the measure of MRPK that a planner who seeks to minimize misallocation cares about. In this subsection, we show that r^{social} is correlated to traditional measures of ARPK, which we view as a form of validation of our measuring framework.

Table 3 reports the results for regressions of the type:

$$\log ARPK_{i,t} = \alpha_{j,t} + \beta \log(r_{i,t}^{social} + \delta) + \epsilon_{i,t} \tag{20}$$

where $ARPK_{i,t}$ is a measure of the ARPK, $\alpha_{j,t}$ are sector-quarter fixed effects (NAICS4, as commonly used in the misallocation literature), and $\epsilon_{i,t}$ is the error term. Each column corresponds to a different measure of ARPK and/or sample. Columns (1)-(2) refer to the Y-14 sample, while (3)-(5) focus on a restricted sample of Y-14 loans that we are able to match to Compustat firms. We consider different measures of ARPK that have been used in the literature: columns (1) and (3) compute ARPK as sales over fixed assets for the Y-14 and Compustat, respectively. Columns (2) and (4) use earnings before interest, taxes and depreciation over fixed assets. Finally, column (5) uses a measure of value-added, only available for Compustat firms.[13] The table shows that the social cost of capital is significantly

[13]Computing value added requires information on the cost of intermediates, which is available neither in the Y-14 nor Compustat. For Compustat firms, it is possible to infer a measure of value added by using the method described in Donangelo et al. (2019): some firms report total annual labor expenses. This can be combined with the number of employees to compute an average wage by industry and year. We then use

correlated with most measures of ARPK, even after controlling for industry and time, with this correlation being stronger and more statistically significant for the VA-based measure in column (5). The remaining lines of the table report the variance of the log of each ARPK measure, and the implied amount of misallocation (on average) that would be obtained by using that respective measure in our sufficient statistic for misallocation. Use of the ARPK measures in the Y-14 data results in extremely large losses from misallocation: 67% and 48% for sales and EBITDA, respectively. Specializing to the Compustat sample results in lower values, but still an order of magnitude above those implied by r^{social}.

Why these differences? One caveat of our measure of misallocation is that it is a measure of capital misallocation only, abstracting from the efficiency in the allocation of other inputs, as well as product market misallocation (i.e., misallocation due to markup dispersion). On the other hand, our measure relies on basic data elements that are typically available in credit registries maintained by financial regulators all around the world, and does not require detailed information on firm financials that is typically needed to compute ARPK, itself just an approximation of MRPK.

6.4 Relation to Text-based Cost of Capital Measures

We compare our measure of the lender discount rate to the firm-level cost-of-capital measures in Gormsen and Huber (2023). Their data use earnings-call transcripts and other firm information to estimate the cost of capital perceived by publicly traded firms. This object is measured very differently from our loan-level estimate of ρ, but both are intended to capture variation in the cost of capital across firms and over time.[14]

We merge their U.S. firm-quarter data to the Y-14 through Compustat and estimate

$$\widehat{r}_{i,t}^{GH} = \alpha_{j,t} + \beta \rho_{i,t} + \epsilon_{i,t}, \tag{21}$$

the average wage to impute total labor costs to other firms in the industry in a given year. Value added is then calculated as the sum of EBITDA to labor expenses, implicitly assuming that all non-labor costs are related to intermediates.

[14]We compare the Gormsen-Huber measure to ρ, rather than to r^{firm}, because their object is closer to a required discount rate or perceived cost of capital than to our model-implied private marginal return net of default losses.

Table 3: Relation to ARPK Measures

	(1) log $ARPK$, Sales	(2) log $ARPK$, EBITDA	(3) log $ARPK$, Sales	(4) log $ARPK$, EBITDA	(5) log $ARPK$, VA
$\log(r^{social} + \delta)$	0.13***	0.24***	0.18**	0.11	0.30***
	(0.03)	(0.04)	(0.07)	(0.13)	(0.08)
Observations	60700	58636	4551	4407	3694
Adjusted R^2	0.26	0.21	0.67	0.51	0.59
Sector-Quarter FE	Yes	Yes	Yes	Yes	Yes
Sample	Y-14	Y-14	Compustat	Compustat	Compustat
Var(log $ARPK$)	2.04	1.56	0.20	0.25	0.22
Misalloc., $ARPK$, %	66.62	47.79	5.15	6.33	5.62
Var($\log(r^{social} + \delta)$)	0.04	0.04	0.01	0.01	0.01
Misalloc., $r^{social} + \delta$, %	0.93	0.93	0.27	0.27	0.27

Notes: The table reports estimates of Equation (20) for different measures of ARPK. Columns (1) and (3) compute ARPK as sales divided by fixed assets, columns (2) and (4) compute ARPK as EBITDA divided by fixed assets, and column (5) computes ARPK as value added divided by fixed assets. Columns (1) and (2) use the full Y-14 sample (for which ARPK and r^{social} can be computed), while columns (3) through (5) restrict the sample to loans that can be matched to Compustat firms. Value added is calculated as EBITDA plus imputed labor expenses, following the method of Donangelo et al. (2019). The last four rows report the variance of $\log(ARPK)$ and of $\log(r^{social} + \delta)$, as well as the corresponding implied output loss from misallocation. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

where $\widehat{r}_{i,t}^{GH}$ is the Gormsen-Huber predicted cost of capital, and $\alpha_{j,t}$ denotes sector-quarter fixed effects (sectors are measured using four-digit NAICS codes). The fixed effects are included in different combinations across columns.

Table 4 shows that our estimated lender discount rate is positively related to the Gormsen-Huber cost of capital. The coefficient on ρ is positive and statistically significant in all specifications. The relationship remains after controlling for both sector and quarter fixed effects, which means that the correlation is not driven only by aggregate time-series movements or broad industry differences. This provides a useful validation check: although the two measures are constructed from very different data sources, firms with higher lender discount rates in the Y-14 also tend to have higher text-based predicted costs of capital.

6.5 Cross-Country Comparison

Finally, in Table 5, we compare our results to related values for other countries. Our method for deriving the distribution of r^{social} incorporates the joint distribution of interest rates, default probabilities, expected losses given default, and leverage. Although prior work has not combined all of these variables in the ways necessary to apply our method, we can still

Table 4: Relation to Gormsen-Huber Predicted Cost of Capital

	(1)	(2)	(3)	(4)	(5)
ρ	0.23***	0.19***	0.15***	0.07***	0.08**
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)
Observations	2224	2224	2191	2191	1615
Adjusted R^2	0.06	0.33	0.39	0.63	0.80
Quarter FE	No	Yes	No	Yes	No
Sector FE	No	No	Yes	Yes	No
Sector-Quarter FE	No	No	No	No	Yes

Notes: The table reports estimates of Equation (21). The dependent variable is the predicted cost of capital from Gormsen and Huber (2023), a text-based measure constructed from firms' earnings-call transcripts. The sample is the subset of Y-14 loans that we are able to match to the Gormsen-Huber data through Compustat. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

learn something from more commonly reported statistics. We focus on papers that report summary statistics for a representative sample of bank loans, or moneylenders in the case of Aleem (1990). We are able to find four such papers that report the mean and standard deviation of interest rates as well as the mean probability of default. For comparison, we also use our data to provide the same statistics for the United States.[15] Relative to less developed countries, the United States stands out for a low mean and low standard deviation of interest rates. However there is significant heterogeneity among the results for other countries, which does not appear to simply track the level of development. Bank loans in Pakistan and Mexico have a standard deviation of interest rates that is only modestly higher than in the United States, while bank loans in Brazil and loans from moneylenders in Pakistan have an extremely high standard deviation of interest rates.

We can attempt to use these statistics to compute misallocation, at the cost of strong assumptions. We use recovery rates from the World Bank's Doing Business database to

[15]Cavalcanti et al. (2024) report the spread relative to deposit rates; we compute the mean interest rate by adding the mean deposit rate, which they report in the paper, to the mean spread. Beraldi (2025) reports the spread relative to the Mexican 28-day interbank rate; we compute the mean interbank rate over this time period and add this to the mean spread. For both papers, we use the distribution of spreads to get the standard deviation of the spread. Beraldi (2025) reports quantiles instead of the standard deviation, so we use the 90/10 percentile range and a normal approximation to infer the standard deviation. We convert reported nominal rates to real rates by subtracting average annual inflation in the respective country during the period of analysis.

provide information on expected losses given default. Since we do not have information on firm leverage, we use the lender's cost of capital, ρ, in place of the social cost of capital, r^{social}. We then use the fixed rate formula for ρ and assume that the probability of default and the losses given default do not vary across firms. This allows us to compute a cost of misallocation, which we show in the last row. The cost of misallocation we compute for the United states is similar to the actual cost that we computed earlier, although this is no guarantee that the same is true for other countries.

There are two main takeaways from our misallocation analysis. First, the United States seems to have the most efficient credit markets of the countries in our table: Misallocation is moderately higher for bank loans in Mexico and moneylenders in Pakistan, and substantially higher in Brazil and among Pakistani banks. However the relationship between development and credit market efficiency varies across settings, even for countries like Brazil and Mexico that are at similar levels of economic development. We derived these results using readily available data but at the cost of adding additional strong assumptions. This encourage future work to estimate the sufficient statistic formula using the micro data of credit registries across different countries.

Table 5: Cross-Country Comparison of Cost of Capital

	Aleem	Khwaja and Mian	Cavalcanti et al.	Beraldi	This Paper
Country	Pakistan	Pakistan	Brazil	Mexico	United States
Years of Data	1980–1981	1996–2002	2006–2016	2003–2022	2014–2025
Mean Real Rate (%)	66.8	8.0	83.0	12.4	2.5
SD Real Rate (%)	38.1	2.9	93.3	5.2	1.3
Mean Default Probability (%)	2.7	16.9	4.0	8.9	1.4
Mean Recovery Rate (%)	42.8	42.8	18.2	63.9	66.7
Implied Misallocation (%)	6.5	13.5	21.5	2.8	0.6

Notes: Data for Pakistan are from Aleem (1990) and Khwaja and Mian (2005). Brazilian data are from Cavalcanti et al. (2024), Mexican data are from Beraldi (2025), and U.S. data are from the present paper. Recovery rates are country-level estimates obtained from the World Bank's Doing Business database. Aleem (1990) studies loans from moneylenders, while the other papers study bank loans. We calculate the implied misallocation using the lender's discount rate ρ, rather than r^{social}, because information on firm leverage is not available. To compute the mean and variance of ρ we use the fixed-real-rate formula, and assume that default probabilities and recovery rates are constant across loans within each setting.

7 Conclusion

This paper develops a novel methodology to estimate the cost of capital using credit registry microdata, and examines the implications of dispersion in the cost of capital for misallocation. We show, in a dynamic corporate finance model, the connection between the lender's cost of capital, the firm's cost of capital, and the social cost of capital, and how to measure these objects in the data. We also show how the mean and variance of the social cost of capital can be used as sufficient statistics to measure the output losses from misallocation that arise from credit market imperfections.

After developing this general methodology, we apply it to credit registry data for the United States. We find that although the cost of capital varies across firms, the resulting misallocation is modest in normal times, resulting in output losses of about 0.7%. However, dispersion in the social cost of capital among newly originated loans rose dramatically during the COVID-19 pandemic, driven by a rise in the dispersion of lender discount rates. Moreover, comparing the distribution of the cost of capital in the United States to the distribution in other economies, especially less developed economies, will help us better understand how financial markets contribute to development.

References

Aleem, I. (1990). Imperfect information, screening, and the costs of informal lending: a study of a rural credit market in pakistan. *The World Bank Economic Review 4*(3), 329–349.

Banerjee, A. V. and E. Duflo (2005). Chapter 7 growth theory through the lens of development economics. In *Handbook of Economic Growth*, pp. 473–552. Elsevier.

BCBS (2001, January). The Internal Ratings-Based Approach. Consultative document, Basel Committee on Banking Supervision.

Begenau, J., S. Bigio, J. Majerovitz, and M. Vieyra (2026, None). A q-theory of banks. *The Review of Economic Studies 93*(1), 106–143.

Beraldi, F. (2025, January). Banking relationships and loan pricing disconnect. Job Market Paper, Yale University.

Carrillo, P., D. Donaldson, D. Pomeranz, and M. Singhal (2023, jun). Misallocation in firm production: A nonparametric analysis using procurement lotteries. Technical report.

Cavalcanti, T. V., J. P. Kaboski, B. S. Martins, and C. Santos (2024). Dispersion in financing costs and development. Technical report, National Bureau of Economic Research.

Chiavari, A., J. Faltermeier, and S. Singh Goraya (2025). The return on capital in disaggregated economies: theory and measurement.

Cooley, T. F. and V. Quadrini (2001, December). Financial markets and firm dynamics. *American Economic Review 91*(5), 1286–1310.

Daruich, D. and J. Kozlowski (2023, July). Macroeconomic implications of uniform pricing. *American Economic Journal: Macroeconomics 15*(3), 64–108.

David, J. M., L. Schmid, and D. Zeke (2022). Risk-adjusted capital allocation and misallocation. *Journal of Financial Economics 145*(3), 684–705.

Dickerson, A., C. Julliard, and P. Mueller (2023). The co-pricing factor zoo. Research paper, UNSW Business School.

Dickerson, A., C. Robotti, and G. Rossetti (2026). The corporate bond factor replication crisis.

Donangelo, A., F. Gourio, M. Kehrig, and M. Palacios (2019, None). The cross-section of labor leverage and equity returns. *Journal of Financial Economics 132*(2), 497–518.

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy 81*(3), 607–636.

Faria-e-Castro, M., S. Jordan-Wood, and J. Kozlowski (2024, July). An Empirical Analysis of the Cost of Borrowing. Working Papers 2024-016, Federal Reserve Bank of St. Louis.

Faria-e-Castro, M., P. Paul, and J. M. Sánchez (2024). Evergreening. *Journal of Financial Economics 153*(C).

Federal Reserve Bank of Cleveland (2025). Inflation expectations. https://doi.org/10.26509/frbc-infexp. Accessed: 2025-08-27.

Gilchrist, S., J. W. Sim, and E. Zakrajsek (2013, January). Misallocation and financial market frictions: Some direct evidence from the dispersion in borrowing costs. *Review of Economic Dynamics 16*(1), 159–176.

Gormsen, N. J. and K. Huber (2023, June). *Corporate Discount Rates*.

Gormsen, N. J. and K. Huber (2024, June). *Firms' Perceived Cost of Capital*.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007). The u.s. treasury yield curve: 1961 to the present. *Journal of Monetary Economics 54*(8), 2291–2304.

Haltiwanger, J., R. Kulick, and C. Syverson (2018, jan). Misallocation measures: The distortion that ate the residual. Technical report.

Hopenhayn, H. A. (2014, aug). Firms, misallocation, and aggregate productivity: A review. *Annual Review of Economics 6*(1), 735–770.

Hsieh, C.-T. and P. J. Klenow (2009, November). Misallocation and manufacturing tfp in china and india. *Quarterly Journal of Economics 124*(4), 1403–1448.

Hughes, D. and J. Majerovitz (2025). Measuring misallocation with experiments. Technical report.

Kaymak, B. and I. Schott (2024). Tax heterogeneity and misallocation. *International Finance Discussion Paper* (1393).

Khwaja, A. I. and A. Mian (2005). Do lenders favor politically connected firms? rent provision in an emerging financial market. *The quarterly journal of economics 120*(4), 1371–1411.

Plosser, M. C. and J. A. C. Santos (2018). Banks' Incentives and Inconsistent Risk Models. *The Review of Financial Studies 31*(6), 2080–2112.

Restuccia, D. and R. Rogerson (2008, October). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics 11*(4), 707–720.

Restuccia, D. and R. Rogerson (2017, aug). The causes and costs of misallocation. *Journal of Economic Perspectives 31*(3), 151–174.

Shorrocks, A. F. (2012, January). Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *The Journal of Economic Inequality 11*(1), 99–126.

Appendix

A Proofs

Derivation of Equation (3). We combine the first-order conditions for capital and for debt. Recall the recursive formulation of the model.

$$V(k,b,z) = \max_{k',b'} \pi(k,b,z,k',b') + \beta\mathbb{E}\left[\max\left\{V(k',b',z'),0\right\} \mid z\right]$$
$$\pi(k,b,z,k',b') = f(k,z) + (1-\delta)k - k' - \theta b + Q(k',b',z)(b' - (1-\theta)b)$$

The firm's maximization yields the first-order conditions for tomorrow's capital, k', and tomorrow's debt, b'.

$$0 = \frac{\partial\pi(k,b,z,k',b')}{\partial k'} + \beta\mathbb{E}\left[\mathcal{P}(k',b',z')\frac{\partial}{\partial k'}V(k',b',z') \mid z\right]$$
$$0 = \frac{\partial\pi(k,b,z,k',b')}{\partial b'} + \beta\mathbb{E}\left[\mathcal{P}(k',b',z')\frac{\partial}{\partial b'}V(k',b',z') \mid z\right]$$

We next use the Envelope Theorem to note that $\frac{\partial V(k',b',z')}{\partial k'} = \frac{\partial}{\partial k'}\pi(k',b',z',k'',b'')$ and similarly to note that $\frac{\partial V(k',b',z')}{\partial b'} = \frac{\partial}{\partial b'}\pi(k',b',z',k'',b'')$. Our first-order conditions become:

$$0 = \frac{\partial\pi(k,b,z,k',b')}{\partial k'} + \beta\mathbb{E}\left[\mathcal{P}(k',b',z')\frac{\partial}{\partial k'}\pi(k',b',z',k'',b'') \mid z\right]$$
$$0 = \frac{\partial\pi(k,b,z,k',b')}{\partial b'} + \beta\mathbb{E}\left[\mathcal{P}(k',b',z')\frac{\partial}{\partial b'}\pi(k',b',z',k'',b'') \mid z\right]$$

Next, we take derivatives of the profit function to plug into our first-order conditions. We have:

$$\frac{\partial\pi(k,b,z,k',b')}{\partial k} = f_k(k,z) + (1-\delta)$$
$$\frac{\partial\pi(k,b,z,k',b')}{\partial b} = -\theta - (1-\theta)Q(k',b',z)$$
$$\frac{\partial\pi(k,b,z,k',b')}{\partial k'} = -1 + \frac{\partial Q(k',b',z)}{\partial k'}(b' - (1-\theta)b)$$
$$\frac{\partial\pi(k,b,z,k',b')}{\partial b'} = Q(k',b',z) + \frac{\partial Q(k',b',z)}{\partial b'}(b' - (1-\theta)b)$$

Plugging these expressions in, our first-order conditions now become:

$$0 = -1 + \frac{\partial Q\left(k', b', z\right)}{\partial k'}\left(b' - (1-\theta)\, b\right) + \beta \mathbb{E}\left[\mathcal{P}\left(k', b', z'\right)\left(f_k\left(k', z'\right) + (1-\delta)\right) \mid z\right]$$

$$0 = Q\left(k', b', z\right) + \frac{\partial Q\left(k', b', z\right)}{\partial b'}\left(b' - (1-\theta)\, b\right) + \beta \mathbb{E}\left[\mathcal{P}\left(k', b', z'\right)\left(-\theta - (1-\theta)\, Q\left(k'', b'', z'\right)\right) \mid z\right]$$

We next combine these two first-order conditions. Rather than thinking about investment that is financed through earnings, we want to instead imagine that the firm is financing a marginal unit of capital through borrowing. To do this, we multiply the first-order condition for debt by

$$-\frac{1 - \frac{\partial Q(k', b', z)}{\partial k'}\left(b' - (1-\theta)\, b\right)}{Q\left(k', b', z\right) + \frac{\partial Q(k', b', z)}{\partial b'}\left(b' - (1-\theta)\, b\right)}$$

which reflects the amount of new debt needed to finance a marginal unit of capital. The denominator reflects the amount raised by selling a unit of debt, $Q\left(k', b', z\right)$, plus an adjustment factor, $\frac{\partial Q}{\partial b'}\left(b' - (1-\theta)\, b\right)$, that reflects how the change in the price of debt affects the cost of borrowing. Similarly, the numerator reflects how an increase in capital is partly self-financing, because it lowers the cost of borrowing.

Combining the two equations then yields:

$$\begin{aligned}
\beta \mathbb{E}\left[\mathcal{P}\left(k', b', z'\right)\left(f_k\left(k', z'\right) + (1-\delta)\right) \mid z\right] = & \frac{1 - \frac{\partial Q(k', b', z)}{\partial k'}\left(b' - (1-\theta)\, b\right)}{Q\left(k', b', z\right) + \frac{\partial Q(k', b', z)}{\partial b'}\left(b' - (1-\theta)\, b\right)} \\
& \times \beta \mathbb{E}\left[\mathcal{P}\left(k', b', z'\right)\left(\theta + (1-\theta)\, Q\left(k'', b'', z'\right)\right) \mid z\right]
\end{aligned}$$

Further manipulation then yields:

$$\begin{aligned}
\mathbb{E}\left[\mathcal{P}\left(k', b', z'\right)\left(f_k\left(k', z'\right) + (1-\delta)\right) \mid z\right] = & \frac{1 - \frac{\partial Q(k', b', z)}{\partial k'}\left(b' - (1-\theta)\, b\right)}{1 + \frac{\partial Q(k', b', z)}{\partial b'} \cdot \left(b' - (1-\theta)\, b\right) / Q\left(k', b', z\right)} \\
& \times \mathbb{E}\left[\mathcal{P}\left(k', b', z'\right) \frac{\theta + (1-\theta)\, Q\left(k'', b'', z'\right)}{Q\left(k', b', z\right)} \mid z\right] \\
= & \mathcal{M} \cdot \left(1 + r^{firm}\right)
\end{aligned}$$

where $\mathcal{M}$ is given by the following formula:

$$\begin{aligned}
\mathcal{M} &= \frac{1-\frac{\partial Q}{\partial k'}\left(b'-(1-\theta)\,b\right)}{1+\frac{\partial Q}{\partial b'}\cdot\left(b'-(1-\theta)\,b\right)/Q} \\
&= \frac{1-\frac{\partial\log Q}{\partial\log k'}\cdot\frac{Q}{k'}\left(b'-(1-\theta)\,b\right)}{1+\frac{\partial\log Q}{\partial\log b'}\cdot\frac{Q}{b'}\left(b'-(1-\theta)\,b\right)/Q} \\
&= \frac{1-\frac{\partial\log Q}{\partial\log k'}\cdot\frac{Q\cdot b'}{k'}\frac{(b'-(1-\theta)b)}{b'}}{1+\frac{\partial\log Q}{\partial\log b'}\cdot\frac{(b'-(1-\theta)b)}{b'}} \\
&= \frac{1-\gamma\cdot\frac{Q\cdot b'}{k'}\cdot\frac{\partial\log Q}{\partial\log k'}}{1+\gamma\cdot\frac{\partial\log Q}{\partial\log b'}}
\end{aligned}$$

where $\gamma := \frac{(b'-(1-\theta)b)}{b'}$. This completes the proof. □

Proof of Proposition 2.

$$1=\sum_{t=1}^{T}\left(\frac{P}{1+\rho}\right)^{t}\left[r+\frac{(1-P)}{P}\left(1-LGD\right)\right]+\left(\frac{P}{1+\rho}\right)^{T}$$

Let $x=\frac{P}{1+\rho}$ so

$$1=\left(r+\frac{1-P}{P}\left(1-LGD\right)\right)\frac{x}{1-x}\left(1-x^{T}\right)+x^{T}$$

Guess that $1+\rho=(1+r)\,P+(1-P)\,(1-LGD)$

$$\frac{1-x}{x}=\frac{1}{x}-1=\frac{(1+r)\,P+(1-P)\,(1-LGD)}{P}-1=r+\frac{1-P}{P}\left(1-LGD\right)$$

And, therefore

$$1=1\left(1-x^{T}\right)+x^{T}$$

which validates the guess. □

Proof of Proposition 3. Rearranging Equations (2) and (4), we have

$$
\begin{aligned}
1+\rho &= \frac{\mathbb{E}\left[\mathcal{P}_{t+1}\left(\theta+(1-\theta)Q_{t+1}\right)+\left(1-\mathcal{P}_{t+1}\right)\frac{\phi(k_{t+1})}{b_{t+1}}\middle|\, k_{t+1},b_{t+1},z_t\right]}{Q_t} \\
&= 1+r_t^{firm}+\mathbb{E}\left[\left(1-\mathcal{P}_{t+1}\right)\frac{\phi(k_{t+1})}{Q_t\cdot b_{t+1}}\middle|\, k_{t+1},b_{t+1},z_t\right] \\
&= 1+r_t^{firm}+(1-P)\cdot(1-LGD)
\end{aligned}
$$

with the last line using our formula for LGD at origination.

□

Proof of Proposition 4. We start with Equation (11), then we use the fact that, by assumption, $\phi'\left(k_{t+1}\right)=\phi\left(k_{t+1}\right)/k_{t+1}$, and then we plug in the definitions of lev and LGD. This yields:

$$
\begin{aligned}
1+r^{social} &= \left(1+r^{firm}\right)\mathcal{M}+(1-P)\cdot\phi'\left(k_{t+1}\right) \\
&= \left(1+r^{firm}\right)\mathcal{M}+(1-P)\cdot\frac{\phi\left(k_{t+1}\right)}{k_{t+1}}\cdot\frac{Q_t b_{t+1}}{Q_t b_{t+1}} \\
&= \left(1+r^{firm}\right)\mathcal{M}+(1-P)\cdot(1-LGD)\cdot lev
\end{aligned}
$$

Plugging in our formula for r^{firm} from Proposition 3 yields

$$
\begin{aligned}
1+r^{social} &= \left(1+\rho-(1-P)\left(1-LGD\right)\right)\mathcal{M}+(1-P)\cdot(1-LGD)\cdot lev \\
&= (1+\rho)\,\mathcal{M}+(lev-\mathcal{M})\cdot(1-P)\left(1-LGD\right)
\end{aligned}
$$

□

B Data

B.1 Details on Data Cleaning and Construction

While the FR Y-14Q Schedule H.1 data goes back to 2011, we keep only data from 2014Q4 due to data quality and consistency of reporting issues.

Borrowers. We drop all loans to borrowers without a Tax Identification Number. We keep only Commercial & Industrial loans to nonfinancial U.S. addresses, i.e. lines reported on FR Y-9C equal to 3, 4, 8, 9, and 10. We drop all borrowers with NAICS codes 52 (Finance and Insurance), 92 (Public Administration), 5312 (Offices of Real Estate Agents and Brokers), and 551111 (Offices of Bank Holding Companies), as some financial companies are classified under the later two NAICS codes in our sample.

Loans. We drop all loans with a negative committed exposure, or for which the utilized exposure exceeds the committed exposure as these are likely to be mistakes. We drop all observations for which the origination date exceeds the current date, and all those for which the maturity date precedes the current date.

We keep only "vanilla" term loans (Facility type equal to 7), and we thus exclude Type A, B, and C term loans, as well as bridge term loans. We keep only loans that are classified as fixed or variable rate, and drop mixed interest rate variability loans. We keep only loans with maturity between 1 and 30 years, thus excluding very short-term and very long-term loans. We keep only loans with interest rates in the 1st-99th percentiles for fixed rate loans, and spread in the 1st-99th percentiles for variable rate loans, as some of the very high and low rates/spreads are likely to be data errors. Additionally, we drop loans with interest rates higher than 50% at origination. We also drop loans for which the probability of default and the loss given default are either missing or outside of the $[0, 1]$ intervals. We also drop loans for which the probability of default is equal to 1, as that is an indicator that the loan is in default.

B.2 Estimating $\mathcal{M}$

In this section, we argue that our calibration of $\mathcal{M} = 1$ is a good approximation. To this end, we provide estimates of this object in the data. Recall that this object was defined as

$$\mathcal{M}_t := \frac{1 - \gamma \times \frac{Qb'}{k'} \times \frac{\partial \log Q}{\partial \log k'}}{1 + \gamma \times \times \frac{\partial \log Q}{\partial \log b'}}$$

Given estimates for the function Q, γ, and firm leverage Qb'/k' we can compute $\mathcal{M}$ for every observation (loan origination) in our data. The main challenge is to estimate Q as a function of firm borrowing and investment. This function can either be obtained by solving a calibrated version of our model, or estimated non-parametrically in the data. In this subsection, we present results for the latter approach.

First, we compute Q for every loan origination in the data. In a model setting such as ours, where loans are modeled as perpetuities that decay at a geometric rate θ, we can write Q as the present value of all future payments, discounted at the contractual interest rate r:

$$Q = \frac{\theta + (1-\theta)Q}{1+r} = \frac{\theta}{r+\theta}$$

r is directly observed in the data, and we can apply the common approximation that θ is equal to the inverse of the loan maturity, $\theta = 1/T$. This allows us to compute Q for every loan origination in the data.[16]

The model establishes that Q is a function of firm investment k', firm borrowing b', as well as the current level of productivity z. Additionally, Q should also depend on the lender's cost of capital ρ. We therefore approximate (the log of) Q as a polynomial of these four variables. We measure firm investment as (the log of) tangible assets at loan origination, firm borrowing as (the log of) total debt owed by the firm at loan origination, firm productivity as the (the log of) sales over tangible assets (a measure of TFPR following Hsieh and Klenow (2009)).

[16]We measure r as the real interest rate at origination, computed as the nominal contractual rate minus average expected inflation over the maturity of the loan.

The lender's cost of capital ρ is measured as in the main text. We therefore estimate:

$$
\begin{aligned}
\log Q_i = \alpha &+ \beta_k \log k_i + \beta_b \log b_i + \beta_z \log z_i + \beta_\rho \rho_i \\
&+\beta_{k,k}(\log k_i)^2 + \beta_{k,b} \log k_i \times \log b_i + \beta_{k,z} \log k_i \times \log z_i + \beta_{k,\rho} \log k_i \times \rho_i \\
&+\beta_{b,b}(\log b_i)^2 + \beta_{b,z} \log b_i \times \log z_i + \beta_{b,\rho} \log b_i \times \rho_i \\
&+\beta_{z,z}(\log z_i)^2 + \beta_{z,\rho} \log z_i \times \rho_i + \beta_{\rho,\rho}(\rho_i)^2 + \epsilon_i
\end{aligned}
$$

The resulting estimates can be used to compute the partial derivatives of $\log Q$ with respect to investment and borrowing. Qb'/k' is measured in a consistent manner, as the sum of total liabilities plus new borrowings divided by total assets plus new borrowings. Finally, we take advantage of the fact that at the steady state, $\gamma = \theta = 1/T$.

Figure 7 presents the histogram for the estimated $\mathcal{M}_i$ in our sample. Clearly, the distribution is extremely concentrated around 1. Figure 8 replicates our measure of misallocation, when computed accounting for heterogeneity in $\mathcal{M}$, and compares it to our baseline, showing that the two measures are extremely similar, both in terms of magnitudes and dynamics. Taken together, these results suggest that our assumption that $\mathcal{M} = 1$ is a good one.

B.3 Cross-sectional Heterogeneity

We also explore the correlation between the cost of capital and firm-level covariates. We regress $\log(1 + r)$ separately on log leverage, log return on assets, and log assets. We conduct this analysis for interest rates, ρ, r^{firm}, and r^{social}. The results are shown in Table 6. Of the three covariates, the best predictor is the return on assets. Return on assets is positively associated with each outcome in the separate regressions. In the joint specification, this association remains positive for ρ, r^{firm}, and r^{social}, but becomes negative for the real contractual interest rate. We interpret these relationships as descriptive rather than causal. The adjusted R^2 for the cost-of-capital regressions ranges from 0.19 to 0.39, reflecting the explanatory power of the covariates together with industry and time fixed effects.

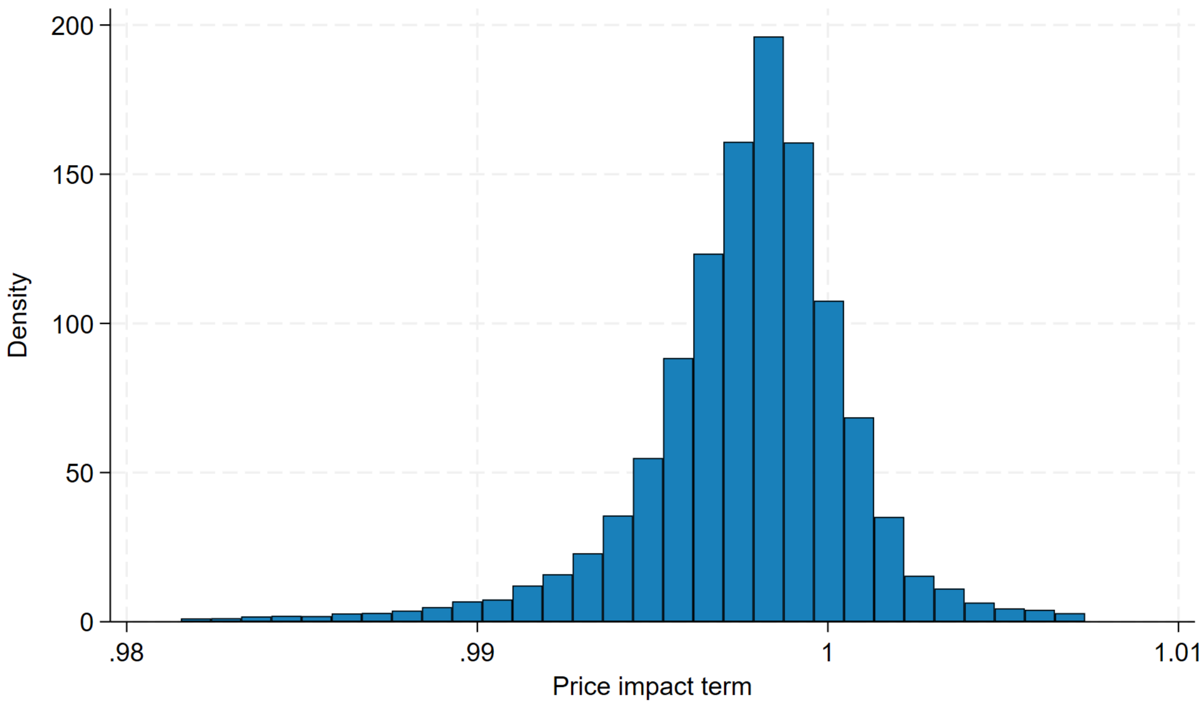
Table 6: Determinants of Capital Costs and Spreads

Panel A: Real Contractual Rate				
	(1)	(2)	(3)	(4)
log leverage	0.099***			0.079***
	(0.00)			(0.00)
log roa		0.014***		-0.054***
		(0.00)		(0.00)
log assets			-0.182***	-0.188***
			(0.00)	(0.00)
Observations	64693	61396	62055	61389
Adj. R2	0.48	0.47	0.50	0.50
NAICS4, Quarter FE	yes	yes	yes	yes
Panel B: Lender Discount Rate				
	(1)	(2)	(3)	(4)
log leverage	0.008**			0.016***
	(0.00)			(0.00)
log roa		0.040***		0.028***
		(0.00)		(0.00)
log assets			-0.047***	-0.033***
			(0.00)	(0.00)
Observations	64693	61396	62055	61389
Adj. R2	0.39	0.38	0.38	0.38
NAICS4, Quarter FE	yes	yes	yes	yes
Panel C: Firm's Cost of Capital				
	(1)	(2)	(3)	(4)
log leverage	-0.075***			-0.058***
	(0.00)			(0.00)
log roa		0.058***		0.097***
		(0.00)		(0.00)
log assets			0.078***	0.106***
			(0.00)	(0.00)
Observations	64693	61396	62055	61389
Adj. R2	0.21	0.19	0.20	0.21
NAICS4, Quarter FE	yes	yes	yes	yes
Panel D: Social Cost of Capital				
	(1)	(2)	(3)	(4)
log leverage	0.107***			0.118***
	(0.00)			(0.00)
log roa		0.068***		0.064***
		(0.00)		(0.00)
log assets			-0.029***	0.020***
			(0.00)	(0.00)
Observations	64693	61396	62055	61389
Adj. R2	0.32	0.30	0.30	0.31
NAICS4, Quarter FE	yes	yes	yes	yes

Standardized beta coefficients; Standard errors in parentheses

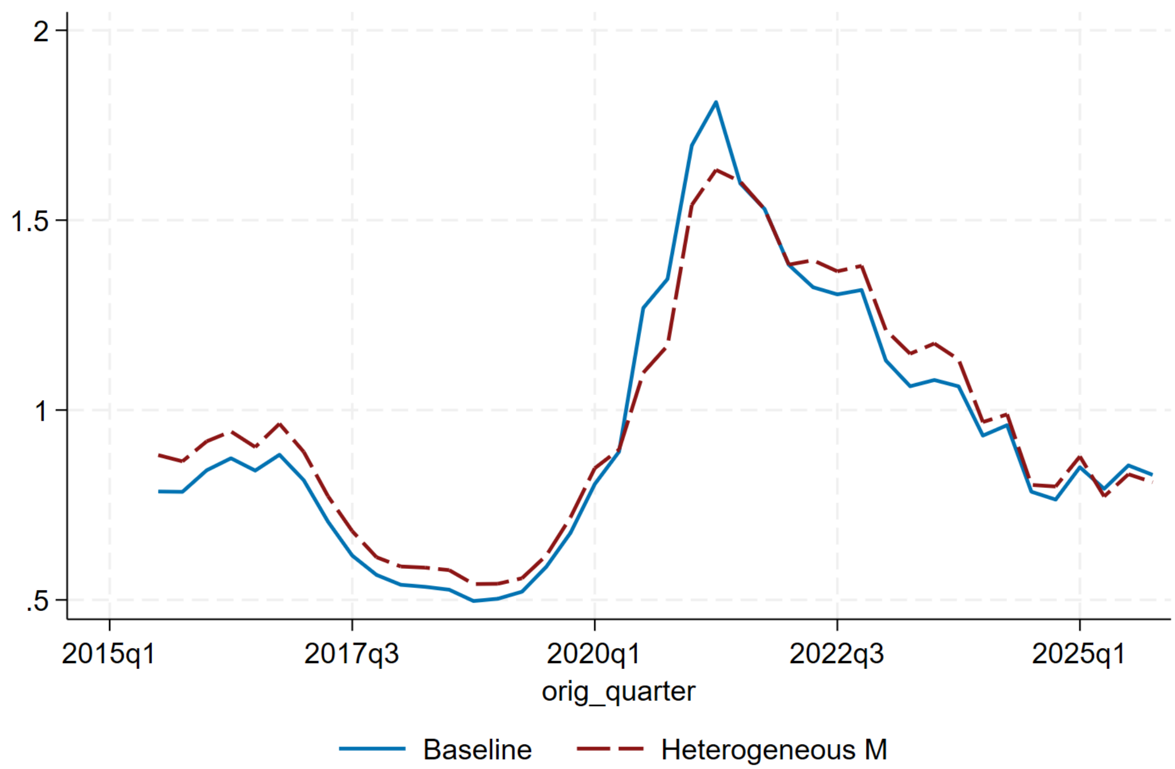
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 7: Histogram for Estimated $\mathcal{M}_i$



Notes: The histogram plots the distribution across loans of the estimated price feedback multiplier, $\mathcal{M}_i$. The estimates are constructed using the procedure described in Appendix B.2.

Figure 8: Implied Misallocation with $\mathcal{M} = 1$ vs. Estimated $\mathcal{M}_i$.

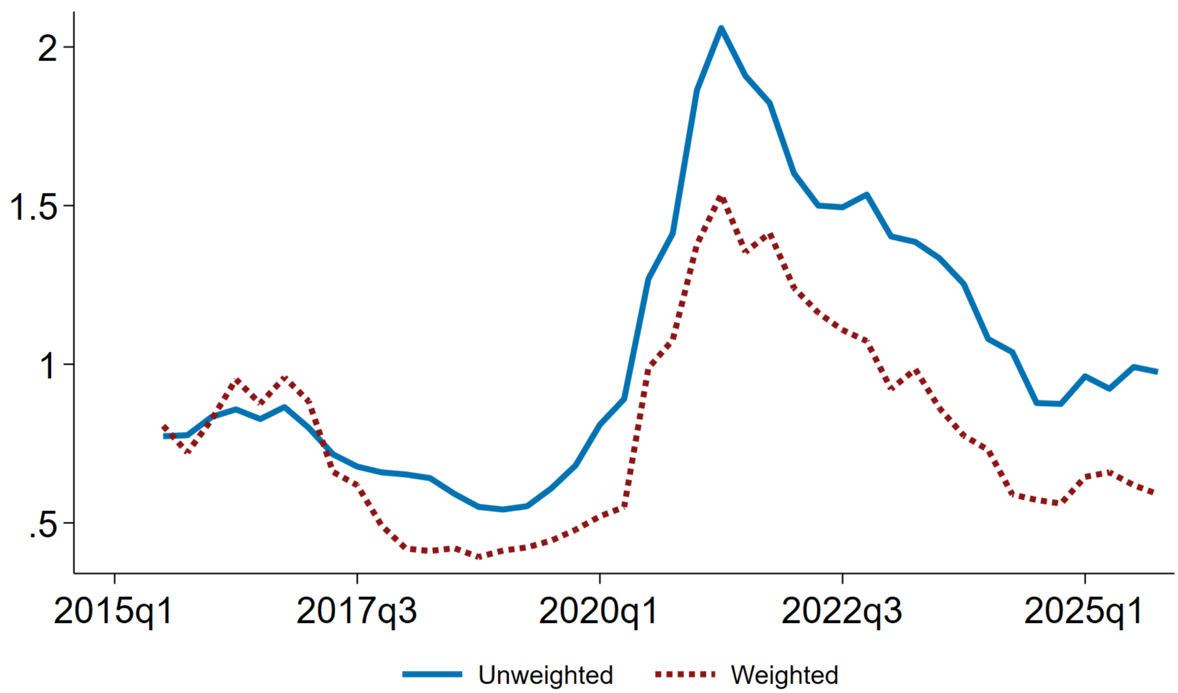


Notes: The figure compares the baseline measure of misallocation, which sets $\mathcal{M}_\rangle = 1$, with the measure constructed using loan-specific estimates of $\mathcal{M}_\rangle$. The estimates of $\mathcal{M}_\rangle$ are constructed using the procedure described in Appendix B.2. Both series are reported as four-quarter moving averages.

B.4 Misallocation Weighted by Loan Size

This subsection examines whether the baseline misallocation series is driven by small loans. In the main analysis, each loan receives equal weight within a quarter. Here, we recompute the sufficient statistic weighting each origination by utilized exposure, so that larger loans receive more weight in the mean and dispersion of the social cost of capital. Figure 9 shows that the weighted and unweighted series are similar before the pandemic, but the weighted measure rises less during 2020–2021 and falls more sharply afterward. Thus, the increase in misallocation is not driven by the largest loans.

Figure 9: Misallocation: Unweighted vs. Weighted by Loan Size

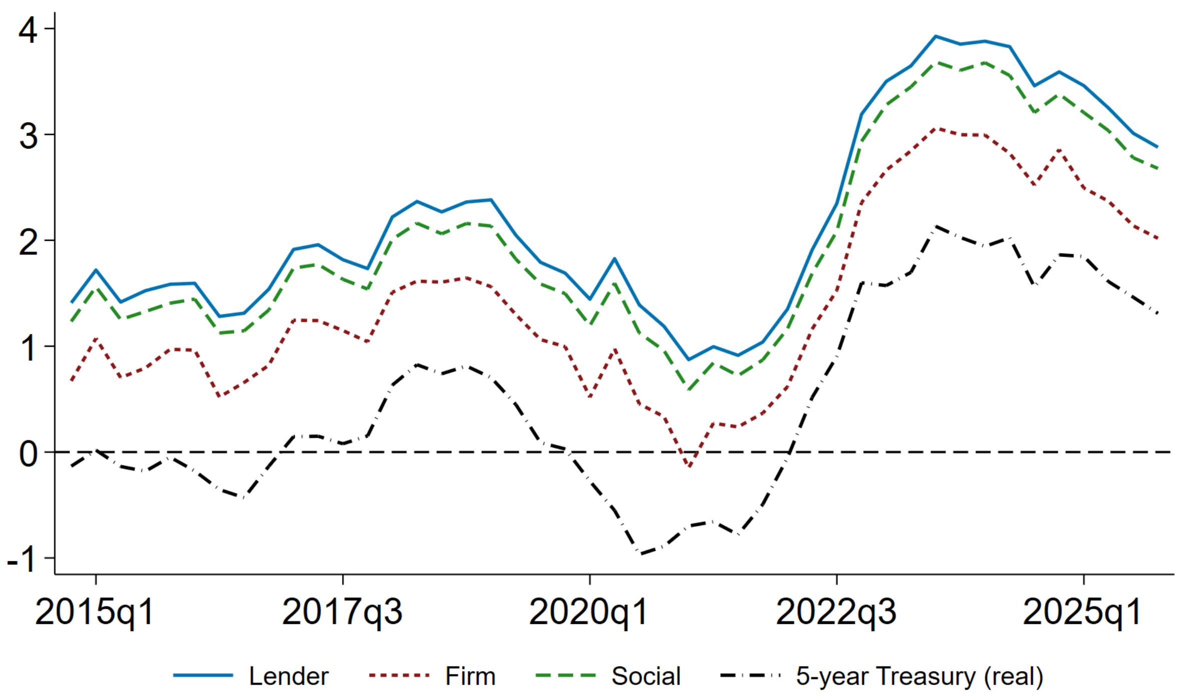


Notes: The figure compares our baseline measure of misallocation, which uses unweighted moments of the distribution of r^{social}, with a measure that weights each loan origination by loan size, which we measure as the utilized exposure of the loan. Both series are reported as four-quarter moving averages.

B.5 Robustness: Results for Five-Year Loans

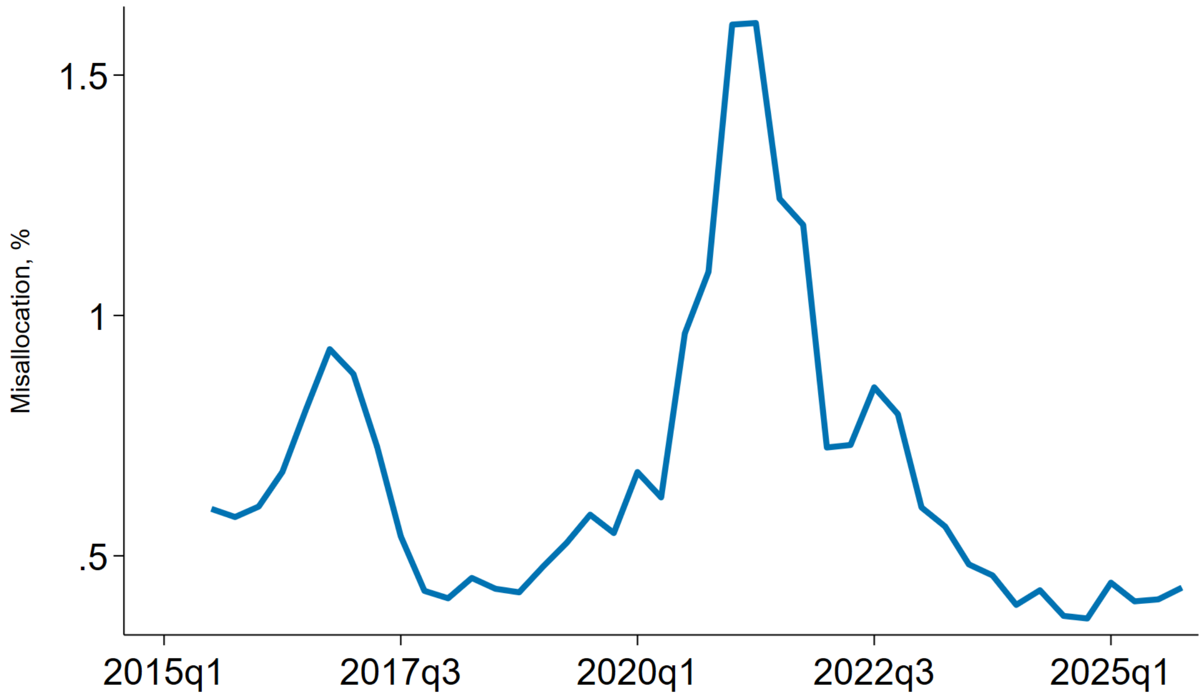
This subsection repeats the main empirical analysis on the subset of loans with maturities close to five years. This restriction addresses the concern that differences in loan maturity, or the mapping from expected future rates into current loan prices, may drive the cross-sectional dispersion in the estimated costs of capital. Five-year loans are also useful because they can be compared directly to the real five-year Treasury rate. The average lender, firm, and social costs of capital in this restricted sample follow the same broad time-series pattern as in the full sample, and the implied misallocation series remains modest in normal times but rises sharply around 2020–2021. The results therefore suggest that the main findings are not an artifact of pooling loans with different maturities.

Figure 10: Cost of Capital for Five-Year Loans



Notes: This figure plots the average cost of capital by quarter of origination, for the sample of loans with a maturity of five years. The real five-year Treasury rate is plotted for comparison. All rates are annual and reported in percent.

Figure 11: Misallocation for Five-Year Loans



Notes: This figure plots the implied cost of misallocation by quarter of origination, for the sample of loans with a maturity of five years. The series is reported as a four-quarter moving average.

C Risk Adjustment

This appendix describes the approach used to construct the risk-adjusted misallocation measure in Section 6.2. The exercise uses traded corporate bonds to measure factor exposure, stock returns to estimate market prices of risk, and Y-14 firm characteristics to impute factor loadings for borrowers that do not have traded securities. The bond-return and factor inputs are based on the Open Bond Asset Pricing data and code pipeline (Dickerson et al., 2026).

Let b index a traded corporate bond, p index a traded stock, i index a firm, m index months, and t index quarters. Let $j \in \{M, H, S\}$ index the market excess return factor, high book-to-market stock returns minus low book-to-market stock returns factor, and small-cap stock returns minus large-cap stock returns factors, respectively.

Step 1: Estimate Bond Betas. We first estimate bond-level factor loadings using duration-adjusted bond returns. For each bond CUSIP b, we estimate

$$r^{du}_{b,m} = \alpha^B_b + \beta^B_{b,M} F_{M,m} + \beta^B_{b,H} F_{H,m} + \beta^B_{b,S} F_{S,m} + u^B_{b,m}, \tag{22}$$

where $r^{du}_{b,m}$ is the bond return net of the duration-matched Treasury return and $F_{j,m}$ are the Fama-French factors. The estimated loadings are denoted $\widehat{\beta}^B_{b,j}$.

Step 2: Estimate Stock Betas. We next estimate stock-level factor loadings. For each stock p, we estimate

$$r^S_{p,m} - r^f_m = \alpha^S_p + \beta^S_{p,M} F_{M,m} + \beta^S_{p,H} F_{H,m} + \beta^S_{p,S} F_{S,m} + u^S_{p,m}. \tag{23}$$

where $r^S_{p,m}$ is the stock return and r^f_m is the risk-free rate. The estimated loadings are denoted $\widehat{\beta}^S_{p,j}$.

Step 3: Estimate Factor Prices. Next, we want to estimate the factor prices. We can do this either using the cross section of bonds or the cross section of stocks. We choose to use stocks, as the sample size is much larger and more likely to be held by a broad set of investors, and because equity factors play an important role in pricing duration-adjusted corporate bond returns (Dickerson et al., 2023). Using the estimated stock betas, we estimate monthly Fama-MacBeth factor prices from cross-sectional stock regressions:

$$r^S_{p,m} - r^f_m = \gamma_{0,m} + \gamma_{M,m}\widehat{\beta}^S_{p,M} + \gamma_{H,m}\widehat{\beta}^S_{p,H} + \gamma_{S,m}\widehat{\beta}^S_{p,S} + \eta_{p,m}. \tag{24}$$

We average the monthly estimates within each quarter to obtain $\widehat{\gamma}_{j,t}$. In the risk-adjustment exercise, we use a three-year moving average of the quarterly factor prices, as in Chiavari et al. (2025):

$$\overline{\gamma}_{j,t} = \frac{1}{12}\sum_{\ell=0}^{11} \widehat{\gamma}_{j,t-\ell}. \tag{25}$$

Step 4: Predict Bond Betas for Compustat Firms. For firms with both stock betas and bond betas, we estimate a factor-specific mapping from stock betas to bond betas:

$$\widehat{\beta}^{B}_{i,j} = a_j + b_j \widehat{\beta}^{S}_{i,j} + e_{i,j}. \tag{26}$$

We estimate this mapping at the bond level. Each bond CUSIP enters as a separate, equally weighted observation, so firms with multiple traded bonds contribute multiple observations. We then collapse the fitted values to the issuer level using an unweighted mean. Because the fitted value depends only on the issuer's stock beta, it is identical across bonds issued by the same firm.

We then use the fitted value

$$\widetilde{\beta}^{B}_{i,j} = \widehat{a}_j + \widehat{b}_j \widehat{\beta}^{S}_{i,j} \tag{27}$$

as the predicted bond beta for Compustat firms with traded equity.

Step 5: Merge Predicted Bond Betas with Y-14. The predicted Compustat-firm bond betas are merged to Y-14 borrowers. This produces a matched sample with predicted bond betas and Y-14 firm characteristics.

Step 6: Predict Bond Betas for All Y-14 Firms. On the matched Y-14 sample, we project the predicted bond betas on firm characteristics observed in the Y-14. For each factor j, we estimate

$$\begin{aligned}\widetilde{\beta}^{B}_{i,j} = {} & c_{0,j} + c_{1,j}\log(\text{assets}_{i,t}) + c_{2,j}\log(\text{sales}_{i,t}) + c_{3,j}\log(\text{cash}_{i,t}) \\ & + c_{4,j}\text{leverage}_{i,t} + c_{5,j}\text{roa}_{i,t} + \lambda_{\text{NAICS2}(i),j} + \nu_{i,t,j}. \end{aligned} \tag{28}$$

The NAICS2 fixed effects allow factor exposure to vary across broad industries. We then use the fitted values from Equation (28) to predict factor loadings for Y-14 borrowers:

$$\begin{aligned}\widehat{\beta}^{Y14}_{i,t,j} = {} & \widehat{c}_{0,j} + \widehat{c}_{1,j}\log(\text{assets}_{i,t}) + \widehat{c}_{2,j}\log(\text{sales}_{i,t}) + \widehat{c}_{3,j}\log(\text{cash}_{i,t}) \\ & + \widehat{c}_{4,j}\text{leverage}_{i,t} + \widehat{c}_{5,j}\text{roa}_{i,t} + \widehat{\lambda}_{\text{NAICS2}(i),j}. \end{aligned} \tag{29}$$

Step 7: Compute the Risk Adjustment. For each Y-14 borrower-quarter, the aggregate-risk compensation is the inner product of the predicted Y-14 bond betas and the smoothed stock-market factor prices:

$$\mathrm{RA}_{i,t} = \widehat{\beta}^{Y14}_{i,t,M}\overline{\gamma}_{M,t} + \widehat{\beta}^{Y14}_{i,t,H}\overline{\gamma}_{H,t} + \widehat{\beta}^{Y14}_{i,t,S}\overline{\gamma}_{S,t}. \tag{30}$$

The adjustment excludes the Fama-MacBeth intercept and uses only the factor-risk component. We annualize the adjustment as

$$\mathrm{RA}^{ann}_{i,t} = (1 + \mathrm{RA}_{i,t})^{12} - 1. \tag{31}$$

Step 8: Compute Risk-Adjusted Returns and Misallocation. We subtract the annualized risk adjustment from the lender discount rate:

$$\rho^{RA}_{i,t} = \rho_{i,t} - \mathrm{RA}^{ann}_{i,t}. \tag{32}$$

We then recompute the firm return and social return using $\rho^{RA}_{i,t}$ and construct the risk-adjusted misallocation measure.